

SECTION 5

ASSESSING THE PROGRESS OF THE LEARNER

'Lack of assessment and feedback, based on observation of performance in the workplace, is one of the most serious deficiencies in current medical education practice.'

John Norcini and Vanessa Burch 2007

OVERVIEW

The assessing of the learner is arguably the most important task for the teacher. Students can walk away from bad teaching, but they cannot avoid bad assessment.

Chapter 28 Six questions to ask about assessment

An understanding of basic concepts and approaches will help you to do a good job.

Chapter 29 Written and computer-based assessment

Written questions have a role to play alongside other assessment methods.

Chapter 30 Clinical and performance-based assessment

An assessment of a student's clinical competence is key to an assessment of their ability to practise medicine.

Chapter 31 Portfolio assessment

Portfolio assessment is a response to changes in medical education including the emphasis on professionalism and the need to give students more responsibility for their own learning.

Chapter 32 Assessment for admission to medicine and postgraduate training

A range of assessment tools and approaches are available to assist with key selection decisions.

Chapter 33 Evaluating the curriculum

Curriculum evaluation is an essential part of the educational process. The focus in curriculum evaluation is on quality improvement.

Six questions to ask about assessment

An understanding of basic concepts and approaches will help you to do a good job.

The assessment of a student's or trainee's learning is important not only for the student or trainee but also for the teacher, the course organiser, the accrediting body and the public as the consumer. Important decisions are taken about students as a result of the scores they achieve in examinations. Teachers and the other stakeholders need to know if students have achieved the appropriate level of mastery to move on to the next part of their training programme and if, on completion of their training, they are competent to practise as a doctor in a particular context.

Although assessment is important, it is one of the most difficult areas in medical education on which to get agreement. What constitutes a fair examination and what are the criteria for passing a student? Assessment is an area in which there have been significant developments as to what constitutes 'good practice' and these will be highlighted in this chapter.

Assessment procedures have been criticised by students, by professional bodies and by those outside medicine. In a recent court case, a judge criticised a nursing school for a failure to identify in its assessment procedures a nurse who proved grossly incompetent and demonstrated unprofessional behaviour after she qualified. For the student, assessment may be seen as analogous to playing in a cricket match where the rules have not been clearly specified in advance and are constantly being changed by the umpire. Students may perceive the examiner as threatening and as someone whose aim is to catch them out and find fault with them (Figs 28.1 and 28.2).

Problems with assessment are serious: students can walk away from bad teaching but they are unable to do so with assessment if they are to achieve the qualification they seek. That assessment is a key and integral part of curriculum development is often not recognised. Issues relating to assessment should be seen not only as a testing or measurement problem but as inextricably linked to the learning outcomes and teaching methods. Course design and assessment are inseparable.

When thinking about assessment it is useful to think about six questions:

1. Who should assess the student?
2. Why assess the student?
3. What should be assessed?



FIG 28.1 Students may perceive the examiner as threatening.

4. How should the student be assessed?
5. When should the student be assessed?
6. Where should the student be assessed?

It is important to think about the overall programme of assessment, including the tools used and how they are implemented, and not to overemphasise one aspect such as the psychometric properties of the assessment instruments.

WHO SHOULD ASSESS THE STUDENT?

One reason why assessment is complex and the teacher's responsibilities may be unclear is that there is a range of stakeholders involved. These include:

- international accrediting bodies
- national accrediting bodies such as the General Medical Council in the UK and, in the USA, The National Board of Medical Examiners
- professional bodies, for example the Royal Colleges in the UK and the National Boards in the USA
- the public and patients
- the individual school in which the student is enrolled
- the department or course committee responsible for teaching the subject
- the individual teacher
- the students themselves.

In medical schools in the UK, the assessment process is overseen by the General Medical Council (GMC) and the implementation is the responsibility

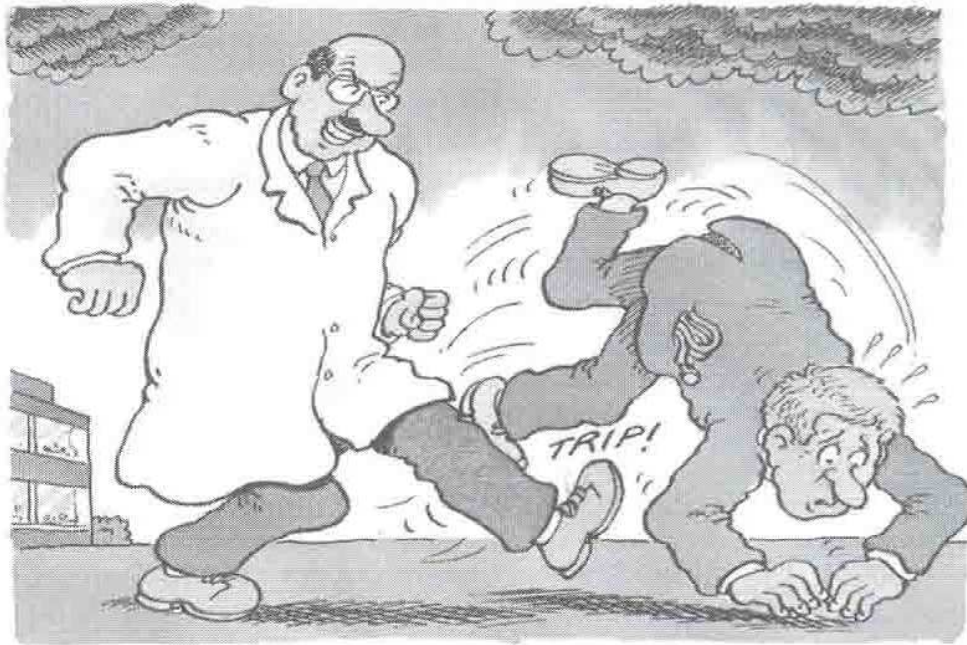


FIG 28.2 Students may perceive the examiner as someone whose aim is to catch them out.

of each medical school. Teachers from other schools serve as external examiners and participate in the development of the school's examinations, their implementation and pass/fail decisions. In contrast, in North America and in some other countries there is a national examination which students are required to pass. Each approach has merits. A national examination, while setting national standards, may stifle innovation in individual medical schools (Harden 2009).

In medical practice, doctors have to take responsibility for assessing their own performance and keeping themselves up-to-date. Not all doctors have the necessary skills or recognise the importance of the responsibility. Students must be prepared for this as part of their undergraduate education with reinforcement throughout their postgraduate training. Problems with unreliability in self-assessment are well recognised and there may also be problems with how students react to the assessment of their own competence. On one occasion, we asked students to mark their own examination paper against a model answer. Some found the procedure so traumatic that they were unable to complete the process and to our surprise required counselling as a result. Training students to become doctors who are inquirers into their own competence, and who are comfortable with this, should be a learning outcome of the curriculum.

Increasing attention is being paid to peer assessment, and the evaluation of students by their peers against certain learning outcomes has become part of some institutions' assessment strategy. This is particularly valuable in the assessment of attitudes where often the student body has a better understanding of individual students' strengths and weaknesses than the teachers.

WHY ASSESS THE STUDENT?

At an early stage in planning an assessment programme it is important to consider the purpose of the assessment. An assessment designed to certify a student's competence to practise as a trainee doctor will be different from the

assessment method used to review a student's progress and provide feedback during course work. Traditionally, assessment has been described as either 'formative', where the main aim is to provide the learners with feedback about their progress, or 'summative' where the aim is to determine whether the learners have achieved the course objectives. This distinction has become blurred with the recognition that summative assessment can also be used to provide feedback to the learner and summative decisions may be based on evidence collected during the training programme.

The purposes that can be served by assessment include:

- **Decisions as to whether the learner is 'fit for purpose'.** Has the learner satisfactorily completed the training programme and achieved the standard expected by the public and professional bodies to practise as a trainee or as a specialist in a particular field of medicine?
- **Assessment of the student's progress during the education or training programme.** It is important to identify deficiencies early in a training programme so that these can be remedied without waiting until a final examination when it will be too late to take the necessary action. This is particularly true with regard to the assessment of behaviour and attitudes.
- **Grading or ranking the student with the aim of identifying the 'best' students among those being assessed.** This 'norm-referenced' approach to assessment is applicable when candidates have to be selected for a limited number of posts, or students selected for admission to medicine where only a set number of places are available. This approach to assessment should not be confused with a 'criterion-referenced' approach where the learner's achievement is assessed against the expected learning outcomes or a set of criteria.
- **Enhancing the student's learning.** Emphasis is placed on 'assessment for learning' rather than 'assessment of learning'. In addition to serving as a tool for accountability, assessment can be a tool to support and improve learning. This is consistent with an 'assessment-to-a-standard' approach where what matters are the standards students achieve rather than the time it takes them to do so.
- **Motivating the student.** It has been demonstrated that assessment has a powerful impact on students and is a major factor in driving their learning. In one medical school we found that, because there was no examination in the subject, students neglected otolaryngology despite the fact that the topic was taught in an imaginative problem-based way. When the subject was routinely included as a station in the final objective structured clinical examination (OSCE), the students' approach to studying the topic changed dramatically.
- **Provision of feedback for the teacher.** The teacher can glean useful information from the student assessment but all too often this source of information is untapped. The analysis of students' scores in one multiple choice question (MCQ) examination revealed that students had performed badly in a question relating to diabetes. This was subsequently found to be related to a weakness in the training programme which had to be addressed.

WHAT SHOULD BE ASSESSED?

A key feature of outcome-based education, as discussed in Section 2, is that assessment is matched closely with the specified learning outcomes. This is referred to as competency-based assessment. What we choose to assess in the education programme demonstrates what we value and many problems encountered with assessment arise from an inadequate consideration of what is being assessed. Assessment drives learning as we have described above. In the absence of a set of learning outcomes what is assessed becomes, for the students, the course objectives. In the past the emphasis in assessment was on the knowledge domain with less attention paid to skills and attitudes. There were a number of reasons for this. Mastery of knowledge was traditionally regarded as of greater importance than the development of attitudes. Knowledge was also easier to assess than other domains and there was a natural tendency to assess what was easy to assess and to shy away from areas where assessment was contentious or difficult. Written assessments, including MCQ papers which tested the knowledge domain, dominated assessment practice. However, someone who can answer correctly a set of MCQs is not necessarily a good doctor and there has been a move to assess in the student or doctor more complex achievement, higher order thinking, clinical skills, attitudes and professionalism.

The introduction of the OSCE stimulated the assessment of psychomotor and other performance related skills, and more recently the adoption of portfolio assessment and multi-source feedback has recognised the importance of the assessment of independent learning and self-assessment skills, attitudes and professionalism.

It is important that, with the many changes advocated in medical education and the different expectations we have of our students in today's curricula, assessment does not lag behind. What we assess must closely match what we expect students to learn.

HOW SHOULD THE STUDENT BE ASSESSED?

A wide range of tools or instruments are now available that can be used to assess the student's competence (Fig. 28.3). Some of these are described in the chapters that follow. Just as important as the assessment tool selected, if not more so, is the way in which the tool is employed. A good tool badly used will yield inappropriate or misleading results. We now have a better understanding of what makes a good assessment:

- **The method should be reliable and consistent.** This relates to the certainty with which a decision can be made about the student's performance on the basis of the test results. For a reliable assessment the measurement instrument must be relatively stable. It would not be good practice, for example, to use an elastic tape measure to measure length. The measurement would not be reliable.

A factor that led to the widespread use of MCQs was that their high reliability could be demonstrated. Problems with reliability associated with tests of clinical competence were highlighted when we found that examiners watching the same clinical performance awarded different scores to the examinees



FIG 28.3 A range of assessment tools are available.

and, watching a video of the students' performance some weeks later, they were not consistent with the marks they had awarded. The problem of reliability was a factor which stimulated us to develop the OSCE.

- **The method should be valid.** In other words the assessment method should measure the learning outcomes intended (Fig. 28.4). The test should be an 'honest' one, testing what it purports to measure. There may be a trade-off at times between reliability and validity. This is illustrated in the classic story of the drunk man who was seen at night looking under a street light for his car keys which he had dropped. When asked why he was looking there when he had dropped them a short distance away, he replied that it was easier to see what was on the ground under the light. His search strategy, although having merit, was not valid in his situation.

Reliability has been emphasised in the past at the expense of validity. What is needed is a test that is both valid and reliable. A test may be reliable but it is of no value if it does not measure what we want to measure. Unfortunately, the more simple a test, the more likely it is to be reliable while at the same time the less likely it is to be valid. Medicine by its nature

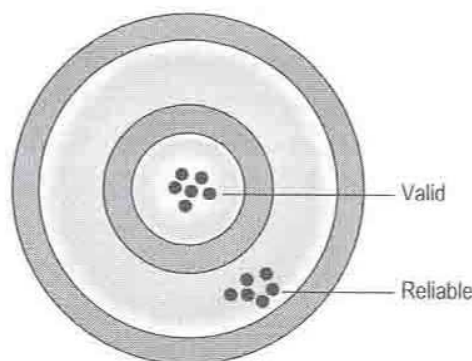


FIG 28.4 Tests should be valid as well as reliable.

is a complex subject and assessment of professionalism or communication skills is by necessity complex if it is to be valid.

- **The method should be feasible** in terms of the resources available and the number of students to be examined. The assessment scheme should not be overly complex and should be capable of being implemented by the teacher in routine practice.
- **The assessment should have a positive impact on the student's learning.** When we first introduced the OSCE in the final examination and included some stations with family physicians as examiners, we found that students spent less time in the library revising their theoretical knowledge and more time learning on the wards and in the community setting. This was in line with the aims of the curriculum.

STANDARD SETTING

A strategy should be adopted which serves as a basis for decisions as to whether a student has reached the standard required to pass the examination. A standard is a statement of whether a student's examination performance is good enough for a particular purpose. Relative or norm-referenced standards, as described above, are based on a comparison of a learner's performance with his or her peers: for example 75% of learners will pass. Absolute standards in contrast are based on the learner's performance in terms of set criteria: for example 60% of the MCQs answered will be correct. Much attention has been focused in recent years on methods of determining the standard expected of students in an examination. Should the pass mark be 60% or should it be 55% or 65% of questions answered correctly? Setting pass standards is a complex task and beyond the scope of this book. Useful introductions to setting standards in examinations can be found in the Exploring Further section at the end of the chapter.

WHEN SHOULD THE STUDENT BE ASSESSED?

Traditionally learners were assessed on their mastery of the subject in a set examination at the end of the course. Increasing emphasis has been placed on collecting evidence about the learner's achievement of the expected learning outcomes during their training or course of study. There are a number of reasons for this. Without the time constraints of a final examination, a much wider sample of the students' performance can be assessed, increasing the reliability of the examination. In addition, the assessment may be more valid in that the assessment tools that can be adopted during the course may assess learning outcomes difficult to assess in an end-of-course final examination. An important additional benefit of in-course assessment is that it provides feedback to the student and teacher and allows time for remediation.

Less frequently undertaken is the assessment of the student at the beginning of the training programme. A number of years ago, we asked third-year students at the beginning of their course in endocrinology to complete the end-of-course examination on day one of the course. The results were surprising. Some students scored less than 10% while other students scored almost

50%, the pass mark for the examination. This provided important evidence as to the need for a course in endocrinology which could be tailored to suit the abilities of the different students. Independent learning modules were developed which were used successfully to replace some lectures on the topic. It could be argued that our first student assessment is when we select students to enter medical studies, or select graduates for a postgraduate training programme (see Chapter 32).

WHERE SHOULD THE STUDENT BE ASSESSED?

Students have been assessed traditionally using written papers in an examination hall environment, or with formal clinical examinations in the hospital ward. With moves towards greater authenticity in assessment, and with assessment seen more as a continuous process, increased attention is being focused on assessment in the work place and in the wide range of contexts where teaching takes place, whether in the hospital or community environment. This is discussed further in Chapter 30.

REACT AND REFLECT

1. Assessment is frequently one of the last things considered when a curriculum is planned. It is often the first thing the learner thinks about. Have you spent enough time considering the assessment of learners in your programme?
2. From the list of purposes given above, what purposes are served by the examinations in your programme?
3. How would you rate your examination against the criteria given for a 'good assessment'?
4. What learning outcomes are assessed?

EXPLORING FURTHER

IF YOU HAVE A FEW HOURS

Bandaranayake, R.C., 2010. Setting and Maintaining Standards in Multiple Choice Examinations. AMEE Guide No. 37. AMEE, Dundee.

An account of the advantages and disadvantages of the more commonly used methods of setting standards in MCQ examinations.

Harden, R., 2009. Five myths and the case against a European or national licensing examination. *Med. Teach.* 31, 217–220.

Norcini, J., Anderson, B., Bollela, V., 2011. Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 conference. *Med. Teach.* 33, 206–214.

This article identifies the current issues in defining criteria for good assessment, makes recommendations on how to proceed and offers some guidance.

Norcini, J., McKinley, D.W., 2010. Standard setting. In: Dent, J.A., Harden, R.M. (Eds.), *A Practical Guide for Medical Teachers*. Elsevier, London (Chapter 41).

A review of some of the common methods for setting standards, and a framework for evaluating their credibility.

Shumway, J.M., Harden, R.M., 2003. AMEE Medical Education Guide No. 25. The assessment of learning outcomes for the competent and reflective physicians. *Med. Teach.* 25, 569–584.

A description of the assessment tools that can be used to assess the different outcome learning.

IF YOU HAVE MORE TIME

- Downing, S.M., Yudkowsky, R. (Eds.), 2009. *Assessment in Health Professions Education*. Routledge, New York.
A comprehensive text devoted specifically to assessment in the health professions.
- Gronlund, N.E., Waugh, C.K., 2009. *Assessment of Student Achievement*, ninth ed. Pearson Education, Upper Saddle River, New Jersey.
A classic text worthwhile consulting, particularly Chapter 1.

- Hodges, B.D., Ginsburg, S., Cruess, R., et al., 2011. Assessment of professionalism: recommendations from the Ottawa 2010 conference. *Med. Teach.* 33, 354–363.
Different ways of thinking about professionalism that can lead toward a multi-dimensional, multi-paradigmatic approach to assessing professionalism at different levels: individual, inter-personal and societal-institutional.

Written and computer-based assessment

Written questions have a role to play alongside other assessment methods.

WRITTEN ASSESSMENT HAS A ROLE TO PLAY

Written approaches to assessment are well established and have been widely used to assess learners' competence in all spheres of education. Despite the greater emphasis being placed on performance assessment, written approaches still have an important role to play. The use of written examinations has come under scrutiny with regard to their match with the expected learning outcomes and their impact on the learners' behaviour. Alternative written assessment methods have been developed as a response to these challenges and increasingly computers have replaced paper and pencil techniques for the delivery of the assessment.

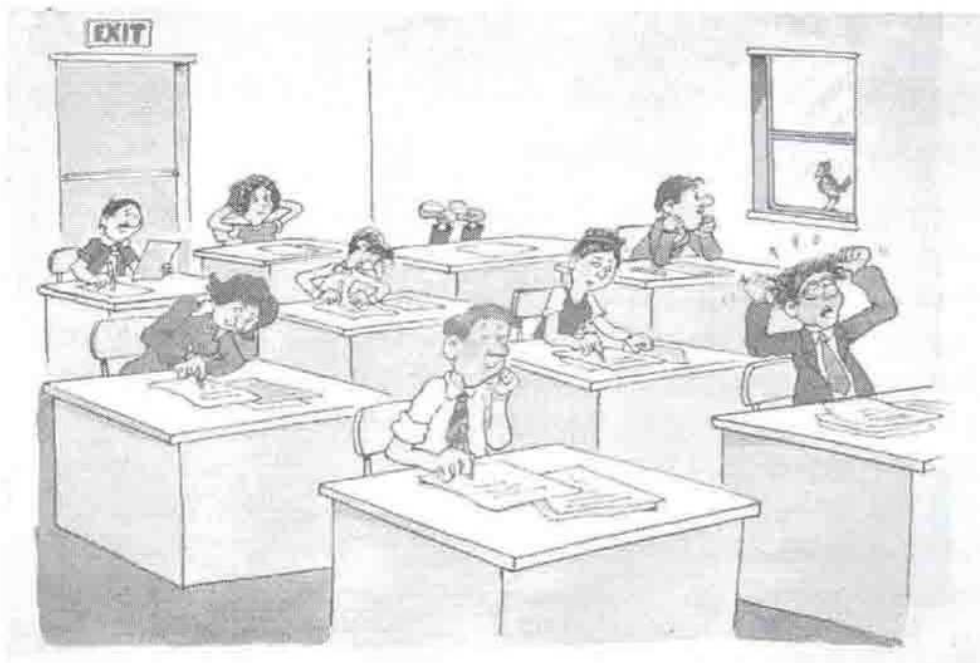


FIG 29.1 A written examination.

THE ELEMENTS IN WRITTEN ASSESSMENT

It is helpful to think of written assessment in terms of the stimulus or task that necessitates a learner's response, the response expected of the learner, the assessment of the student's response and the media or technology used.

THE STIMULUS

The stimulus for the learner's response may be:

- a short statement or question, e.g. 'In which of the following pathologies does a patient typically present with weight loss and an increased appetite?' or 'List the 3 options in the management of a patient with hyperthyroidism.'
- a statement or question with accompanying diagrams or charts, e.g. 'In the diagram which structure is labelled 'A'?'
- a short clinical scenario with a patient's presentation followed by a question, e.g. 'Mrs Wilkie, a 35-year-old waitress complaining of tiredness and nervousness...'
- a more extended patient management problem that develops over a period of time.

THE STUDENT'S RESPONSE

The response expected of the learner can be categorised into:

- Constructed response questions where the student has to write a long or short narrative in response to the stimulus. These include essays, short essay questions and short answer questions.
- Selected response questions where the student has to make a selection from a range of options provided. These include multiple choice questions (MCQs), for example the one best answer or multiple true/false questions and extended matching items.

THE ASSESSMENT OF THE STUDENT'S RESPONSE

The learner's response may be scored:

- Automatically correct or incorrect, as in an MCQ where there is an agreed correct answer. This also applies to short answer questions where the expected answer is limited to a few words. In this case, agreement has to be reached with regard to alternative wording and spelling that is acceptable.
- In a constructed response question, by an examiner based on a holistic impression of the student's response, or with an assigned structured marking scheme.
- In the context of the answers to the questions provided by members of a panel of experts. This is the strategy used in the script concordance test as described below.

STANDARD SETTING

The subject of standard setting was introduced in the previous chapter. A range of methods have been used in a written examination to determine the mark above which students will pass the examination and below which they will fail (Bandaranayake 2008).

It is not uncommon for a pass mark for an MCQ paper to be set at 60–70%. This implies, however, that students need to know only two thirds of the area covered and that it does not matter which third they do not know. Although it is not usual practice, consideration needs to be given to an examination or part of an examination that assesses essential core knowledge, with a pass mark of more than 90% expected for that part of the examination.

TYPES OF WRITTEN ASSESSMENT

ESSAY QUESTION

While the essay remains a common assessment tool in many fields, it is less commonly used in medicine. Provided the question or stimulus is appropriate, the essay can test:

- the learner's general understanding of the topic
- higher level skills including synthesis, organisation of information, analysis, problem solving and evaluation
- written communication skills, a competence often not tested by other written assessment methods although portfolio assessment may be used for this purpose
- aspects relating to attitudes and medical ethics.

Essay questions do have a number of disadvantages as an assessment tool:

- The content area sampled is small compared to an MCQ paper.
- The scoring of the questions is subjective and time consuming.

SHORT ESSAY QUESTION (SEQs)

Short essay questions are designed to sample a wider range of content than the essay question. A student may have to answer twelve 10-minute short essay questions instead of four 30-minute or two 1-hour essay questions. The SEQ has many of the advantages and disadvantages of the essay question.

SHORT ANSWER QUESTIONS (SAQs)

In a short answer question, rather than selecting from a list of choices as in a MCQ, the student is expected to respond to the question in one, two or a limited number of words. The format has the advantages of an MCQ in that a wide range of content can be sampled and has the added advantage that it does not simply test recognition of the correct answer from a list of options. It is also easier to set questions that test core or basic knowledge without having to cue the learner with the responses on a list. MCQs offer the advantage that they can be automatically marked but SAQs share this advantage and it is possible to mark an SAQ using the appropriate computer program, although agreement needs to be reached with regard to the answers that are acceptable. SAQs are now routinely used in examinations such as the progress test at the medical school in Dundee and merit wider consideration and adoption as an alternative to the MCQ.

MULTIPLE CHOICE QUESTIONS (MCQs)

Multiple choice questions can sample objectively a wide range of a student's knowledge and understanding. The downside for the examiner is not the marking of the responses but the setting of the questions. Several question banks have been developed in medicine and these allow schools to share questions.

Many MCQ examinations test mainly recall of knowledge rather than in-depth understanding or application. This has had a detrimental effect on how students study. To counteract this, there has been a move to introduce greater authenticity to the MCQ through the use of clinical scenarios in the stem or stimulus.

Many formats have been described but two approaches have dominated:

- Single best option where the learner is required to select the best response from four or five alternatives.
- Multiple true/false questions where the student has to categorise as true or false each of five statements relating to the stem. Guessing is a more important consideration with this type of question and as a result the marking scheme may be more complicated. There has been a move away from using this type of question although it does test a wider range of knowledge and the questions are easier to set without the need for distracters as in the single best option type of question.

EXTENDING MATCHING QUESTIONS (EMQs)

The EMQ is a type of written question in the selected response category that provides an alternative to the standard MCQ. The EMQ consists of a list of around 20 options relating to a theme or topic. This may be a list of drugs, diseases, laboratory investigations, symptoms, explanations or pathologies. Following the lead-in stem or stimulus, usually in the form of a patient scenario, the student has to select the most appropriate answer from the list of options. There may be one, two or more questions using the same list of options. An advantage is that, with the extended list of options from which the student has to select the answer, the effect of cuing as found in the MCQ is minimised. The questions are less time consuming to produce as several scenarios can use the same list of options. As with MCQs a computer can be used to score the answers.

MODIFIED ESSAY QUESTIONS (MEQs)

The MEQ is a sequence of questions based on a patient scenario that develops over a period of time. After the first question is answered, further information is provided about the patient and this is followed by another question. The typical MEQ may have six or more questions and both SAQs and MCQs may be used. There are difficulties with scoring this format and it is now rarely used.

SCRIPT CONCORDANCE TEST (SCT)

This is a relatively new form of written assessment. Students are given a brief patient scenario and asked to make judgements regarding diagnostic possibilities or management options. To allow decisions to be made that reflect medical practice, the student is given sufficient clinical detail but a certain amount of uncertainty, imprecision or incompleteness is deliberately built into each case in

order to simulate real-life clinical situations. The scoring is based on the amount of agreement between the student and a panel of experts. The approach may be useful, particularly in the assessment of a student's clinical reasoning.

THE TECHNOLOGY

Questions can be posed, responses recorded and examinations scored either on paper or on a computer. Computers are now widely used for MCQs: the students answer on a computer and their responses are automatically scored. There is some interest in adaptive testing where the questions presented to the students on the computer are determined by their responses to earlier questions. A wrong answer to a question may generate additional related questions that allow the student's understanding or lack of understanding of the area to be explored further. This may increase the reliability of the examination by increasing the number of questions used in the areas where the student's understanding is in doubt. This approach is not in routine use at the moment.

REFLECT AND REACT

1. Given the learning outcomes to be tested and the resources available, which format of written test is appropriate to your setting?
2. To what extent does a written examination in your course test more than factual recall?
3. Consider how you can best provide feedback to students about their performance in a written examination. Simply providing a mark, for example 62%, is not sufficient.

EXPLORING FURTHER

IF YOU HAVE A FEW HOURS

Amin, Z., Boulet, J.R., Cook, D.A., et al., 2011. Technology-enabled assessment of health professions education: consensus statement and recommendations from the Ottawa 2010 Conference. *Med. Teach.* 33, 364–369.

This paper highlights the changing nature of ICT in assessment and the challenges that need to be addressed when technology is incorporated into assessment.

Bandaranayake, R.C., 2008. Setting and maintaining standards in multiple choice examinations. *AMEE Guide No. 37. Med. Teach.* 30, 836–845.

An account of the advantages and disadvantages of the more commonly used methods of setting standards in MCQ examinations.

Case, S.M., Swanson, D.B., 2002.

Constructing Written Test Questions for the Basic and Clinical Sciences, third ed. (revised). National Board of Medical Examiners, Philadelphia.

A classic description of how to prepare MCQs.

IF YOU HAVE MORE TIME

Schuwirth, L.W.T., Van der Vleuten, C.P.M., 2010. Written assessment. In: Cantillon, P., Wood, D., (Eds.), *ABC of Learning and Teaching in Medicine*. second ed. (Chapter 9). John Wiley and Sons, Chichester.

A useful overview of the use of written examinations in medicine.

An assessment of a student's clinical competence is key to an assessment of their ability to practise medicine.

THE IMPORTANCE OF CLINICAL ASSESSMENT

Written assessment instruments assess what students know and how they apply this knowledge to a written problem. Clinical and work-based assessment instruments assess the clinical and practical skills of the examinee and how knowledge is applied in the clinical context. The clinical examination is of key importance in the assessment of the learner's competence to practise medicine, and in many schools is the cornerstone in qualifying examinations.

The assessment of competence can be distinguished from performance assessment. Tests of competence, such as the objective structured clinical examination (OSCE), demonstrate in a controlled situation what an examinee is capable of doing. Performance assessment tools such as record analysis or multi-source feedback assess what the individual does in practice. The Miller pyramid provides a framework for assessment with the bottom of the pyramid being the assessment of knowledge and the higher levels of the pyramid being the assessment of performance (Fig. 30.1).



FIG 30.1 The Miller pyramid.

THE PATIENT

Central to the clinical examination is the interaction of the examinee with a patient. The role of the patient in the encounter varies depending upon the type of interaction between the examinee and the patient expected. For the purpose of the examination, the patient may be a 'real' patient, a simulated patient or a computer representation used as a patient substitute.

There are benefits to be gained from the use of a range of patient representations in the clinical examination. The choice of patient representation will be influenced by what is being assessed, the level of standardisation required, the required realism or fidelity and the local logistics including the availability and relative costs associated with the use of real patients and trained simulated patients (Collins and Harden 1998).

'Real' patients

The traditional clinical examination was based on 'real' patients and they are the basis for the patient encounter in a work-based assessment. 'Real' patients are widely used in OSCE stations although this is less common in North America. The examination organisers may have access to a bank of patients with a range of pathologies such as a patient with goitre or a patient with rheumatoid arthritis.

Simulated patients

Difficulties in standardising real patients and a lack of availability in some situations led to the development of simulated or standardised patients. These have been used for assessment as well as teaching. The simulated patient, as described in Chapter 25, is usually a lay person who has undergone various levels of training in order to provide a consistent clinical scenario. The examinee interacts with the simulated patient in the same way as if they were taking a history, examining or counselling a real patient. Simulated patients are used most commonly to assess history taking and communication skills or physical examination where no abnormality is found. Simulated patients have also been used to simulate a range of physical findings including, for example, different neurological presentations. The term 'standardised patient' has been used to indicate that the person has been trained to play the role of the patient consistently and according to specific criteria.

Simulators and models

Simulators, from the very basic models used to assess skills such as skin suturing to the more complex interactive whole-body manikins such as SimMan, have been used increasingly in medical training as outlined in Chapter 25. They have an important role to play in assessment. The Harvey cardiac manikin, for example, has been used at an OSCE station to assess skills in cardiac

auscultation. Simulators are valuable to assess procedural and practical skills including the insertion of intravenous lines, catheterisation and endoscopy technique. While simulators have played a key role in competence assessment in other fields, notably with airline pilots, simulators have been slow to make an impact in assessment in medicine. The situation has changed rapidly and such devices now play a prominent role in clinical assessment. Indeed in some instances surgeons are allowed to perform a procedure in clinical practice only after they have demonstrated competence on a simulator.

Computer-based simulations

The development of virtual patients is another area where rapid progress has been made. While early computer-based simulations were little more than patient management problems delivered through a computer, computer simulations now include high fidelity models of the patient care environment and the 'patient' responds to the management and therapeutic efforts of the examinee.

THE EXAMINER

In addition to the student and the patient, the third key element in clinical or performance assessment is the examiner. The role of the examiner is to collect evidence about the examinee's behaviour in the context of the assessment and to pass judgement on the examinee's competence or performance. The examiner may be a clinician, another healthcare professional or a simulated patient. After appropriate training simulated patients are, in some situations, particularly in North America, used to assess the student's performance in an OSCE. Other members of the healthcare team frequently contribute to multi-source feedback assessment as described below.

Whatever the assessment method used, it is important to include a number of examiners. A problem with the long case in a traditional clinical examination was over-reliance on the ratings of one or two examiners. In contrast, the OSCE has input from a number of examiners which is a major advantage.

A STUDENT'S PROFILE

A wide range of different learning outcomes or competencies including clinical skills, practical procedures, decision making and problem solving, collaboration and team working and professionalism and attitudes is assessed in a clinical examination. It makes little sense simply to allocate a percentage score for each component and then to sum these to produce a total mark of, for example, 62%. Excellence in carrying out an examination or practical procedure should not compensate for unprofessional behaviour or an inappropriate attitude. The answer is not to agonise over the relative importance of each element of competence and the allocation of a percentage for that element but to produce a competence profile for the examinee. This indicates for each candidate, as shown in Figure 30.2, the domains where their performance is satisfactory or perhaps excellent and those domains or competencies where their performance falls short of what is expected.

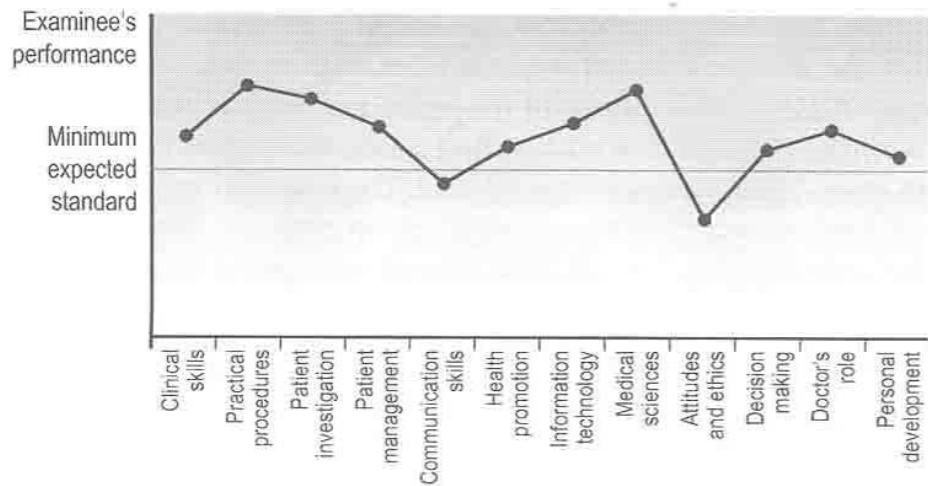


FIG 30.2 An examinee's profile demonstrating competence in some areas but not others.

FEEDBACK

The provision of feedback to the examinee is an important part of a clinical assessment. As discussed in Chapter 2, this should be specific, non-evaluative and timely. It is an essential element of 'assessment for learning' and should guide the learner in their further studies. It is important that when the assessment process is planned, time is allowed for the provision of feedback. The OSCE can be considered as an example of the different approaches that can be adopted:

- Following an OSCE, students' checklists and ratings sheets for each station can be returned to the examinee with the examiner's comments attached.
- Following an OSCE, a meeting with the teacher can be scheduled and the examination is discussed with individuals or the class as a whole.
- During an OSCE, time can be allocated at the end of each station for the provision of feedback from the examiner to the examinee.
- After completion of a station, the examinees review at the following station their own performance at the previous station as recorded on the examiner's checklist and compare this with a model performance as shown on a video recording.

APPROACHES TO CLINICAL AND PERFORMANCE ASSESSMENT

A range of methods have been used to assess the student's or trainee's competence in a controlled clinical environment and to assess how they perform in the work place or real clinical situation.

THE OBJECTIVE STRUCTURED LONG EXAMINATION RECORD (OSLER)

In the traditional long case, the examinee takes a history and examines a patient over a period of up to an hour unobserved by the examiner. Following this the examiner meets with the student and over a 20–30-minute period discusses the patient and the examinee's findings and conclusions. As a replacement for

the 'long case' component of a clinical examination, the OSLER was proposed as a more objective and valid assessment of the student's clinical competence (Gleeson 1997).

Over a 30-minute period the examiner uses a structured score sheet to assess the candidate's performance with a patient in the following areas:

- history taking scored in relation to pace and clarity of presentation, communication skills, systematic approach and the establishment of the case facts
- physical examination rated in relation to a systematic approach, examination technique, and the establishment of the correct physical findings
- the ability to determine appropriate investigations for the patient
- the examinee's views on the management of the patient
- the clinical acumen and overall ability to identify and present a satisfactory approach to tackling the patient's problems.

The examiner grades the examinee's performance for each of the areas assessed, taking into account the difficulty of the case, with a rating of 'excellent', 'very good', 'pass', 'bare pass', 'below pass' or 'seriously below pass'. The examiner also records an overall grade for the complete performance.

THE OBJECTIVE STRUCTURED CLINICAL EXAMINATION (OSCE)

The OSCE was introduced in 1975 in response to criticisms about the reliability and validity of the traditional clinical examination. It has been adopted worldwide and is now recognised as the gold standard for the assessment of clinical competence. Students rotate round a series of stations at a predetermined time interval. Each station focuses on one or more aspects of competence, such as history taking, physical examination or carrying out a procedure.

A typical OSCE lasts 2 hours and has 24 stations with 5 minutes being allocated for each station. This allows a wide sample of competencies to be assessed. Some OSCEs have fewer stations with a longer time allocated for each station. In general, however, it is preferable to use shorter stations rather than longer ones as this increases the reliability and validity of the examination. An OSCE with 24 5-minute stations is preferable to an examination with 12 10-minute stations. Where a task cannot be completed in the 5-minute period there are three options:

1. The task can be specified so that it can be completed in 5 minutes.
2. The task can be spread over two linked stations with the first part of the task, for example assessing a patient's record, undertaken at the first station and the subsequent task of counselling the patient assessed at the following station.
3. A station may be duplicated in the circuit to allow students to spend double the set time at the station. This may be useful for a history taking station.

In the OSCE, any subjective bias attributed to an examiner is reduced as the student will encounter a number of examiners during the course of the examination. What is assessed at each station is agreed in advance and a marking schedule is produced which is completed by the examiner. It is important that examiners are fully briefed and trained in advance.

An OSCE should include preferably 'real', simulated or standardised patients, manikins and simulators as each has advantages in what they can offer for the purposes of the assessment. This is not always possible and 'real' patients or simulated patients may dominate in the OSCE.

The content of the examination together with the competencies to be tested is carefully planned on a blueprint in advance of the OSCE. This ensures that the examination tests a range of competencies in relation to different aspects of medical practice. Examples of the types of stations that can be included in an OSCE are given in Appendix 5.

The OSCE offers major attractions as a reliable and valid test of clinical competence. A major advantage is that the format can be easily adapted for use in a wide range of different settings

MINI CLINICAL EVALUATION EXERCISE (MINI-CEX)

In a Mini-CEX the examinee is engaged in an authentic work place based patient encounter. This may be in the in-patient setting, out-patient department or emergency room. The examiner watches the examinee take a focused history, perform relevant parts of the physical examination, provide a diagnosis and present a management plan. The encounter usually lasts about 15–20 minutes and the examiner spends 5 minutes giving the examinee feedback. The performance is scored on a 6- or 9-part scale ranging from below expectations through borderline and meeting expectations to above expectations. The examinee is responsible for the timing of the encounter and for the selection of the patient. The Mini-CEX is repeated on a number of occasions during a clinical attachment with different examiners and different patients.

The Mini-CEX was developed for use in a postgraduate setting but it has been used also in undergraduate education. In this situation the duration of the encounter is often increased from 20 to 40 minutes.

The Mini-CEX has the advantage that it ensures that the clinical skills of the student or trainee are observed in the work place setting by the clinician or trainer and that feedback is given to the learner. The briefing of the examiner and examinee are important.

DIRECT OBSERVATION OF PROCEDURAL SKILLS (DOPS)

This is a variation of the Mini-CEX designed to assess and give feedback on a student's or trainee's procedural skills. As in the Mini-CEX the student or trainee is observed in the work place carrying out the procedure on 'real' patients. The trainee selects the timing and the procedure to be assessed from a prescribed list including, for example, central venous line insertion, arterial blood sampling, electrocardiography and intubation. The examiner may be a clinician or another member of the healthcare team. As with the Mini-CEX, DOPS was designed for use in postgraduate education but has also been applied in undergraduate education.

CASE-BASED DISCUSSION (CbD)

The CbD is used principally in postgraduate training. The trainee selects several case records in which they have made entries regarding patients that they have seen recently. The examiner selects one patient record and explores aspects of it with the trainee. The assessment is designed to assess application of knowledge, decision making and ethical issues as well as medical record keeping. Dimensions assessed include the trainee's clinical assessment of the patient, investigations and referrals, treatment, follow-up and future planning, professionalism and overall clinical judgement. Each dimension is scored on a 6-point scale. Fifteen minutes is allowed for the examination followed by 5 minutes feedback.

MULTI-SOURCE FEEDBACK (MSF) OR 360 DEGREES EVALUATION

MSF has been used for many years in industry and adopted more recently in medicine. It is used in postgraduate and continuing education to assess the practising doctor. Evidence is systematically collected from a number of individuals who are in a legitimate position to make a judgement about the doctor's performance. The individuals may be senior or junior colleagues, other members of the healthcare team, administrators, patients or students. In this way, different perspectives are brought to bear on the evaluation of the doctor. The individual is asked to complete a structured questionnaire relating to the doctor's performance. A '1 to 5' or a '1 to 7' rating scale can be used and comments may also be recorded. The questions asked may be the same or vary for different groups of respondents. Information is collated so that the ratings remain anonymous and the results are fed back to the doctor. The aim is to provide a fair and balanced view of the doctor's behaviour and abilities, particularly in areas such as communication skills, leadership, team working, punctuality and reliability.

MSF is less frequently used in undergraduate education but is sometimes included for assessment purposes in a student's portfolio. This may include peer assessment of professionalism.

MSF offers many advantages, in particular the assessment of the doctor in the real-life practice context. It also has potential disadvantages, in particular the risk of providing damaging and over-harsh feedback.

REFLECT AND REACT

1. The assessment of clinical competence is a central issue in medical education and whatever your role as a teacher, you should be aware of the different approaches and tools available.
2. You may find yourself responsible for the organisation of an OSCE. The OSCE is a flexible tool. See a couple in action then design one to suit your own needs.
3. You may have to serve as an examiner in an OSCE or be the person responsible for providing students with feedback on their performance. Your role may be to advise students or trainees who are about to sit an

OSCE as to how best to prepare for it. Ensure that you are familiar with the examination details.

4. Performance assessment in the work place is particularly challenging. The examiner and examinee both need to be committed to its success. Think about how you might distinguish between competence and performance in your trainees.

EXPLORING FURTHER

IF YOU HAVE A FEW HOURS

- Boursicot, K., Etheridge, L., Setna, Z., et al., 2011. Performance in assessment: consensus statement and recommendations from the Ottawa conference. *Med. Teach.* 33, 370–383.
The conclusions of a group of experts in the area.
- Boursicot, K.A.M., Roberts, T.E., Burdick, W.P., 2010. Structured assessments of clinical competence. In: Swanwick, T. (Eds.), *Understanding Medical Education: Evidence, Theory and Practice*. Wiley-Blackwell, Chichester, pp. 246–258.
Practical advice on planning, running and scoring an OSCE.
- Collins, J.P., Harden, R.M., 1998. The use of real patients, simulated patients and simulators in clinical examinations. *AMEE Medical Education Guide No. 13.* *Med. Teach.* 20, 508–521.
Patients are a key element in a clinical examination. This is a description of the different ways patients can be represented.
- Gleeson, F., 1997. Assessment of clinical competence using the Objective Structured Long Examination Record (OSLER). *AMEE Medical Education Guide No. 9.* *Med. Teach.* 19, 7–14.
A description of how the long case can be improved in a clinical examination.
- Harden, R.M., Gleeson, F.A., 1979. *ASME Medical Education Guide No. 8.* Assessment of Medical Competence Using an Objective Structured Clinical Examination (OSCE). ASME, Edinburgh.
An early description of the OSCE but still with much relevance today.
- Hill, F., Kendall, K., 2007. Adopting and adapting the mini-CEX as an undergraduate assessment and learning tool. *The Clinical Teacher* 4, 244–248.
How the Mini-CEX can be used in undergraduate education.

Norcini, J., Burch, V., 2007. Work-place Based Assessment as an Educational Tool. *AMEE Medical Education Guide No. 31.* AMEE, Dundee.
A description of the range of tools used in work place based assessment.

IF YOU HAVE MORE TIME

- Gronlund, N.E., Waugh, C.K., 2009. *Assessment of Student Achievement.* Pearson, Upper Saddle River, New Jersey.
Chapter 9 in this classic text provides an overview of performance assessment.
- Holmboe, E.S., Hawkins, R.E., 2008. *Practical Guide to the Evaluation of Clinical Competence.* Mosby, St Louis.
A more detailed account of assessment of clinical competence.
- Norcini, J., Holmboe, E., 2010. Work-based assessment. In: Cantillon, P., Wood, D. (Eds.), *ABC of Learning and Teaching in Medicine.* second ed. Wiley-Blackwell, Chichester (Chapter 11).
An introduction to work-based assessment.
- Tekian, A., McGuire, C.H., McGaghie, W.C., 1999. *Innovative Simulations for Assessing Professional Competence: From Paper-and-Pencil to Virtual Reality.* University of Illinois at Chicago Department of Medical Education.
A collection of essays describing innovative applications of simulation technology in assessment and certification in different disciplines.

Portfolio assessment is a response to changes in medical education including the emphasis on professionalism and the need to give students more responsibility for their own learning.

WHAT IS A PORTFOLIO?

A portfolio is a collection of evidence that learning has taken place. It is cumulative in the sense that it contains work collected over a period of time rather than the snapshot view obtained with the traditional examination. It is important to appreciate that a portfolio is different from a logbook. In a portfolio the learner's experiences are recorded but also included are reflections on the experiences and a description of the further learning that has resulted.

The portfolio is likely to contain both quantitative graded evidence as well as qualitative descriptions. Students actively collect and select the material for their portfolio which will provide the examiner with evidence of their learning. Portfolios may include evidence of the practical procedures carried out by the student, videotapes of their clinical experiences, evaluations of their abilities in written assessments and reports by clinicians on the student's clinical attachments. Multi-source feedback from nurses, other members of the healthcare team and patients may be included. The students individualise their portfolio by selecting the evidence relating to their own personal experience. The evidence included in a portfolio is limited only by the degree of the designer's creativity.

WHY PORTFOLIOS?

As described in the previous chapter, in the 1970s there was a switch of emphasis from an assessment of students' knowledge to an assessment of their clinical skills including history taking, physical examination and practical procedures. Tools such as the objective structured clinical examination (OSCE) were developed for this purpose. The more recent move to outcome-based education with an emphasis on learning outcomes such as attitudes, professionalism, reflection and self-assessment created the need for a tool that provided a more valid assessment in these areas. There was a need too for an assessment tool that counteracted a reductionist approach to assessment and which provided a more holistic and overall assessment of a student's competence.

My (RMH) daughter completed two honours degree courses. One was in mathematics and computing and the other in fashion design. She was in no doubt that the assessment method used in her fashion course – portfolio assessment – was much more searching and accurate, and fairer as an assessment of competence in a professional area, than the more traditional examinations used in the mathematics and computing course. Her fashion portfolio contained evidence of work completed during her studies and her reflections on this. It contained evidence of the technical and inter-personal skills she had acquired during her training and demonstrated her understanding of the theory that underpinned the work. Talking with her, the potential value of a portfolio as an assessment tool in medicine was apparent.

Portfolios, which for many years have been used in the arts, are now having a major impact as an assessment tool in medicine. The use of portfolios for learning and assessment has now become widespread in medicine and other healthcare professions. Portfolios are an authentic learning and assessment tool that relates to the work of a doctor and reflects a holistic and integrated approach to medical practice.

ADVANTAGES

Portfolios offer a number of advantages as an assessment tool:

- Outcomes and competencies can be assessed that other tools have difficulty in reaching. These may be skills necessary for life-long learning such as self-assessment, reflection and the adoption of appropriate learning strategies. They provide a tool to assess attitudes and professionalism.
- Portfolios include evidence collected over a period of time and provide an overall and holistic view of a student's competence.
- With the range of quantitative and qualitative evidence included and the triangulation from the different sources of evidence, portfolios provide a comprehensive and reliable interpretation of a student's achievement.
- Portfolios represent a personalised approach to assessment, focusing on what individual students achieve in contrast to the more standardised approach with instruments such as the OSCE.
- As the students go through their course of studies the portfolio integrates learning and assessment and focuses the student's attention on the learning outcomes expected. The student's reflection on the basic sciences relevant to a documented patient encounter reinforces the application of theory to practice.
- The portfolio reinforces a student-centred approach to the curriculum with students being given greater responsibility for their own learning and assessment.

IMPLEMENTING PORTFOLIO ASSESSMENT IN PRACTICE

As an aid to those wishing to design and implement portfolio assessment in their own institution, the late Miriam Friedman Ben David described 10 steps in the portfolio assessment process:

1. *Define the purpose.* It should be made clear whether the portfolio is to be used for summative or formative decisions and how it relates to other elements in the assessment process.

2. *Determine the competencies to be assessed.* Identifying the competencies to be assessed by the portfolio is part of a systematic approach to assessment that ensures that all of the expected learning outcomes are assessed. In particular, the portfolio is valuable in assessing learning outcomes such as attitudes and professionalism and the higher levels of Miller's pyramid.
3. *Define the portfolio content.* Evidence should be included in the portfolio that demonstrates a student's achievement of the learning outcomes to be assessed. Students should be given guidelines as to the type of evidence that is acceptable, but there should be a certain amount of freedom of choice. The type of evidence that is included, the student's comments on it, and their ability to assess their own competence is a measurement of their understanding of the learning outcomes.

Examples of the type of evidence that might be included are:

- records of procedures carried out by the student
- annotated details of patient encounters with details of how they contributed to the student's achievement of the learning outcomes
- evaluation of the student during clinical clerkships by members of the healthcare team, including nurses and patients
- videotapes of student interaction with patients.

A practising doctor's portfolio will comprise a dossier of evidence demonstrating the doctor's continuing education and practice achievements. The evidence included in a portfolio should indicate the student's or practising doctor's progress over time.

4. *Develop a marking system.* Specific criteria may be set out for each of the learning outcomes to be assessed with the portfolio, and a global rating used for the achievement of each outcome:

- 4 – excellent, distinguished or superior
- 3 – satisfactory, adequate or competent
- 2 – minimal borderline or marginal
- 1 – unsatisfactory or inadequate.

Several assessors should review each portfolio with an assessment committee taking a final decision.

5. *Select and train the examiners.* The choice of examiners will depend on the purpose of the assessment and the learning outcomes to be assessed. The examiners should include a range of staff from the basic sciences and clinical medicine. Examiners with less experience can be paired with more senior examiners. The training of the examiners is essential for the success of the programme. Faculty members often appreciate their participation in portfolio examinations as it allows them to get to know more about the individual student and his or her capabilities.
6. *Plan the examination process and timetable.* It is necessary to set a deadline for portfolio submissions. Students' failure to meet the deadline is itself evidence of a lack of professionalism. Time should be scheduled for examiners to read each portfolio and to meet to discuss them.

An opportunity should be provided for them to meet, possibly in pairs, with each student to allow the student to defend the portfolio. Finally, time needs to be set aside for the examiners to discuss each student's performance in order to come to a final decision as to the student's achievement of the learning outcomes assessed.

7. *Student orientation.* Students should be informed in writing about the portfolio assessment process and what is expected of them. In general, the more information given to students the more positive they are about the portfolio.
8. *Develop guidelines for decisions.* If portfolios are used for summative pass/fail decisions, standards need to be specified so that there is no doubt about what constitutes a pass or fail. A decision needs to be made whether poor performance in relation to one outcome assessed can be compensated by good or excellent performance in another area, or whether areas are not compensatory. The medical school in Dundee adopted the approach that a student cannot compensate for deficits in one domain, such as attitudes or professionalism, by a good performance in another domain. Students must achieve the minimum expected standard in every domain.
9. *Establish reliability and validity evidence.* What constitutes good reliable evidence should be agreed prior to the implementation of the portfolio. The degree of reliability may be determined in a pilot study. For example, should there be two pairs of examiners for each portfolio or one pair? Triangulation of the evidence in the portfolio from different sources will increase the validity of the decision reached and guide the faculty as to the use of the portfolio results.
10. *Evaluate the portfolio assessment.* Students' and examiners' opinions on the strengths and weaknesses of portfolios as an assessment tool should be sought. Students' performance with the portfolio assessment should be compared with their performance in an OSCE or written examination and ultimately with their subsequent performance as a doctor. When problems are identified relating to the professionalism or performance of practising doctors, were the same issues identified in their portfolios as students?

REFLECT AND REACT

1. Review the value of portfolios as a tool for student assessment. Which of the advantages listed are applicable to the programme for which you have a responsibility?
2. What could be included in your student's or trainee's portfolio that would provide evidence of their achievement of the expected learning outcomes?
3. If you are an examiner for a portfolio assessment it is important that you fully understand the assessment process and the marking scheme used.

IF YOU HAVE A FEW HOURS

- Davis, M.H., Friedman Ben-David, M., Harden, R.M., et al., 2001. Portfolio assessment in medical students' final examinations. *Med. Teach.* 23, 357–366. *A description of how portfolios can be used as the basis for a final examination in a medical school.*
- Driessen, E., van Tartwijk, J., van der Vleuten, C., Wass, V., 2007. Portfolios in medical education: why do they meet with mixed success? A systematic review. *Med. Educ.* 41, 1224–1233. *A description of how portfolios can be used for summative and formative purposes.*
- Friedman Ben-David, M., Davis, M.H., Harden, R.M., et al., 2001. Portfolios as a Method of Student Assessment. AMEE Medical Education Guide No. 24. AMEE, Dundee. *Practical advice on the use of portfolios for the purpose of assessment.*
- van Tartwijk, J., Driessen, E.W., 2010. Portfolios for Assessment and Learning. AMEE Guide No. 45. AMEE, Dundee. *Practical advice on the use of portfolios.*

IF YOU HAVE MORE TIME

- Buckley, S., Coleman, J., Davidson, I., et al., 2009. The Educational Effects of Portfolios on Undergraduate Student Learning: a Best Evidence Medical Education (BEME) systematic review. AMEE, Dundee. *A summary of the evidence for the educational effects of the use of portfolios in undergraduate education.*
- Gronlund, N.E., Waugh, C.K., 2009. Assessment of Student Achievement, ninth ed. Pearson, Upper Saddle River, New Jersey. *Chapter 11 describes the value of portfolios as a means of assessment and how they can be scored.*
- Tochel, C., Haig, A., Hesketh, A., Cadzow, A., et al., 2009. The effectiveness of portfolios for post-graduate assessment and education. BEME Guide No. 12. *Med. Teach.* 31, 299–318. *A systematic review of the evidence on the effectiveness of portfolios across postgraduate healthcare.*

Assessment for admission to medicine and postgraduate training

A range of assessment tools and approaches are available to assist with key selection decisions.

THE IMPORTANCE OF SELECTION

In recent years the methods adopted to select students for admission to medical studies from the large number of applicants have been the focus of attention. There are good reasons for this:

- Owing to low attrition rates, the admission of students to medical studies is almost equivalent to graduating a student. Once students are selected for admission to medicine they will almost certainly complete their medical studies and graduate as a doctor.
- There are political and other pressures to widen access to medical studies and there is wide recognition that doctors should be matched to the community they serve. Some ethnic and social classes may have been disadvantaged in the selection process and this needs to be taken into consideration in the admissions process.
- Once students graduate they are more likely to practise in the geographical area or type of community in which they originally lived.
- The criteria used for selection were based traditionally on academic qualifications. It was assumed that if a student achieved top grades in their studies at school they would automatically develop the competencies expected of a good doctor. This is not necessarily true, and the personal qualities of the potential doctor are recognised to be important as well as their academic qualifications.
- There may be advantages in choosing students whose career goals match the mission of the medical school, for example in relation to a commitment to rural practice.

APPROACHES TO SELECTION FOR ENTRY TO MEDICINE

GRADUATE OR DIRECT FROM SCHOOL ENTRY

In North America, following the Flexner 1910 Report, students were admitted to medical school as graduates following completion of a college course in another subject. In other parts of the world students can enter medical studies

direct from school. There has been a move in some medical schools in the UK, Australia and other countries to graduate-only entry. Other schools have maintained a mixed approach accepting both graduates and school leavers. The issue is a controversial one and the arguments for the different approaches are complex. There is no good evidence to indicate that one approach is preferable to the other, but opting to admit students straight from school does have significant financial advantages and there is no evidence that the product is less satisfactory.

A RANGE OF METHODS

Different approaches have been promoted in the selection of students for admission to medicine. Until recently the emphasis has been on academic ability and intellectual achievement. Attention has been paid to selecting an individual who will do well in the medical school examinations rather than to selecting an individual who will become a competent and good doctor. The ability to perform well in tests may simply predict later test scores and not performance as a doctor. The ability to answer MCQs correctly does not necessarily indicate that a student will be a good doctor. This does not mean that a choice has to be made between someone who has a good academic record and someone who will become a good doctor. They are not mutually exclusive.

In general, the criteria for good assessment discussed in Chapter 28 also apply to the selection process. This must be seen to be fair and to be reliable and valid. For this to be so, evidence to support the selection decision should be obtained from a range of sources. The motivation of the student to study medicine must also be taken in to account. Whatever approach to selection is adopted it is important that criteria such as academic ability, personal qualities and ethnicity and any weighting given to the different factors are made explicit. When the entry requirements are considered, attention should be paid to the expected exit learning outcomes of the medical school. Evidence that the student meets minimal requirements relating to each outcome domain prior to entry to medicine is advisable. For example, a required level of competence relating to the communication or attitudinal domains may be specified prior to admission. Learning outcome frameworks that can be used for this purpose are discussed in Chapter 8.

PERFORMANCE AT SCHOOL

Academic achievement and performance at school, as evidenced by A-level achievements in the UK or grade point averages (GPAs) in the USA, have played a major role in decisions about selection. Correlations have been shown between such indicators and students' performance in medical school, particularly in examinations in the early years.

APTITUDE TESTS

Various aptitude tests have been designed to measure a student's ability to develop skills or acquire knowledge. Most have a knowledge component. Examples are the North America College Admission Test (MCAT), the Graduate Australian Medical Admission Test (GAMAT) and the United Kingdom Clinical Aptitude Test (UKCAT).

AUTOBIOGRAPHICAL NARRATIVE

The application form completed by the student may contain useful information about the applicant that will inform the admission decision. The student may be required to provide an autobiographical narrative to justify their motivation to study medicine. Such forms, however, are at risk of fraud and may be written by a third party.

REFERENCES

References may be sought from the student's school, previous employers or individuals with whom the student has been associated. Such letters of recommendation seek to identify personal qualities such as honesty and application. These references are open to bias and tend to be unreliable and ineffective. Laws in relation to disclosure and removal of confidentiality cast additional doubt on their value.

INTERVIEWS

Interviews have been widely used to complement assessment of the student's achievement and aptitude. While questions have been raised about their reliability, they are used to provide evidence that the student has the attributes expected of a future doctor. Structured or semi-structured interviews, where the questions asked are standardised for the candidates and a rating scale is used, have been shown to be more reliable. Interviews, however, are labour intensive and also subject to bias. In one school, when the measurements were taken for refurbishment of the bars in the student union, it was found that the students admitted to the medical school were taller than average. A possible explanation was that the individual in charge of the interview admission process at that time was 6' 6". When interviews are used, training of the interviewer is important.

THE MULTIPLE MINI-INTERVIEW (MMI)

In recent years, there has been increasing emphasis on more objective assessment of students for the purposes of selection with the use of an OSCE approach. The multiple mini-interview (MMI) consists of a series of 5–8-minute testing stations. A rater at each station assesses the student's performance in areas such as ethical decision making, effective communication, empathy, manual dexterity, knowledge of the healthcare system and critical thinking. The MMI requires fewer examiner hours than the traditional type of interview that involved a panel of examiners. The evidence to date is that the MMI has more predictive validity. It provides also a stimulus to prompt admission committees to be more explicit about the qualities they are looking for in applicants.

SELECTION FOR ADMISSION TO SPECIALTY TRAINING

Many of the issues concerning the selection of students to medical studies also apply to the selection of doctors for postgraduate training. Undergraduate performance and interviews play a significant part in the process. Dean's letters and other references are often used but again their reliability is suspect.

Not much attention has been paid to matching doctors on graduation with the career for which they are best suited. The skills expected of a surgeon will differ in some respects from the abilities and characteristics expected of a doctor working in public health.

REFLECT AND REACT

1. You should familiarise yourself with the selection procedure in your institution. To what extent does this reflect the attributes expected of a future doctor?
2. To what extent do the principles of good assessment described in Chapter 28 apply to the selection process in your institution?
3. If you are involved in a selection committee, an interview panel or as an assessor in an MMI you should ensure that you are familiar with the rules and complete the necessary training.
4. Consider how social accountability of universities and training organisations should be reflected in the selection process and the widening of access.

EXPLORING FURTHER

IF YOU HAVE A FEW HOURS

Eva, K.W., Rosenfield, J., Reiter, H.I., et al., 2004. An admissions OSCE: the multiple mini-interview. *Med. Educ.* 38, 314–326.

A description of this relatively new approach.

Harris, S., Owen, C., 2007. Discerning quality: using the multiple mini-interview in student selection for the Australian National University Medical School. *Med. Educ.* 41, 234–241.

A description of the use of the MMI in the selection of students in one school.

McGaghie, W.C., 2002. Student selection. In: Norman, G.R., van der Vleuten, C.P.M., Newble, D.I. (Eds.), *International Handbook of Research in Medical Education*. Kluwer Academic Publishers, London (Chapter 10). *A useful summary of medical student selection criteria and some of the methods adopted.*

Prideaux, D., Roberts, C., Eva, K., et al., 2011. Assessment for selection for the health care professions and specialty training: consensus statement and recommendations from the Ottawa 2010 conference. *Med. Teach.* 33, 215–223. *Recommendations from an expert group on assessment for selection in medicine.*

Curriculum evaluation is an essential part of the educational process. The focus in curriculum evaluation is on quality improvement.

THE IMPORTANCE OF EVALUATION

The earlier chapters in this section are concerned with the evaluation of the student. In this chapter we look at the evaluation of the medical school or educational organisation and the curriculum in the institution. This is important for a number of reasons:

- With the current emphasis on quality assurance it is important to ascertain whether the curriculum and the teaching are fit for purpose.
- Curriculum development is an ongoing iterative process, which requires to be informed by an evaluation of current practice and with feedback as to problems identified.
- There is a move to recognise excellence in education in medical schools as well as excellence in research.

LEVELS OF EVALUATION

An evaluation of the education programme can be conceptualised at different levels:

- At the macro level the mission of the medical school and the extent to which the school is achieving its overall aims should be examined.
- At the micro level the elements of the curriculum should be evaluated. These include the learning outcomes, the teaching, learning and assessment methods, the educational strategies, the education environment, the management of the curriculum and the engagement with students and staff.

ASSESSMENT OF THE MEDICAL SCHOOL/INSTITUTION

The educational programme within the institution and the management structure should be evaluated as described below. It is important also to take a broader look at the institution and its achievements. Are the aims as set out in the mission of the institution appropriate and are they being achieved?

Curriculum evaluation is an essential part of the educational process. The focus in curriculum evaluation is on quality improvement.

THE IMPORTANCE OF EVALUATION

The earlier chapters in this section are concerned with the evaluation of the student. In this chapter we look at the evaluation of the medical school or educational organisation and the curriculum in the institution. This is important for a number of reasons:

- With the current emphasis on quality assurance it is important to ascertain whether the curriculum and the teaching are fit for purpose.
- Curriculum development is an ongoing iterative process, which requires to be informed by an evaluation of current practice and with feedback as to problems identified.
- There is a move to recognise excellence in education in medical schools as well as excellence in research.

LEVELS OF EVALUATION

An evaluation of the education programme can be conceptualised at different levels:

- At the macro level the mission of the medical school and the extent to which the school is achieving its overall aims should be examined.
- At the micro level the elements of the curriculum should be evaluated. These include the learning outcomes, the teaching, learning and assessment methods, the educational strategies, the education environment, the management of the curriculum and the engagement with students and staff.

scope, people, in, model

ASSESSMENT OF THE MEDICAL SCHOOL/INSTITUTION

The educational programme within the institution and the management structure should be evaluated as described below. It is important also to take a broader look at the institution and its achievements. Are the aims as set out in the mission of the institution appropriate and are they being achieved?

THE MISSION OF THE SCHOOL

An aspect of the mission of a medical school that has attracted attention is the social responsibility and accountability of the school. A school should not be an ivory tower but should be regarded as excellent only if significant attention is paid to its social accountability and how it relates to the community that it serves. Boelen and Woollard (2009) have argued that for a school to be socially responsible it needs to have a commitment to respond to the priority health needs of citizens and society. The impact of the educational institution on society and the public good should be part of the assessment of its success. It has been shown that the most research-active medical schools demonstrate the least social responsibility. Some schools see as their mission the development of future leaders or researchers.

THE SCHOOL'S GRADUATES

Another factor in the assessment of a medical school is the consideration and attention paid to the continuum of medical education and the further professional development of its graduates. Has the school collaborated in the development of its curriculum with the range of stakeholders, including patients and those in the health service responsible for employing the doctors?

INTERNATIONAL DIMENSION

With increased globalisation in medicine and medical care, the extent to which the medical school is adequately equipping its graduates as citizens of the world should be evaluated. Graduates should have a sound knowledge of global issues, the skills for working in an international context and the values of a global citizen.

CURRICULUM EVALUATION

DEFINITION

Curriculum evaluation can be defined as 'a deliberate act of enquiry which sets out with the intention of allowing people concerned with an educational event to make rigorous, informed judgements and decisions about it so that appropriate development may be facilitated' (Coles and Grant 1985).

AIMS OF CURRICULUM EVALUATION

The evaluation of a curriculum can have different purposes:

- Demonstration of the achievement of the minimum standards expected by an accrediting body such as the General Medical Council in the UK, the Liaison Committee on Medical Education (LCME) in the USA or the Australian Medical Council (AMC) in Australia. The World Federation for Medical Education (WFME) has set out minimum standards for educational programmes at undergraduate, postgraduate and continuing education levels.

- Establishing that the programme as set out ('the planned curriculum') is in fact happening ('the received curriculum'). There is often a gap between what is set out on paper and what happens in practice.
- As an essential requirement for curriculum development. Without an evaluation of the curriculum there can be no informed curriculum development. Curriculum evaluation should be concerned not simply with a measurement of the success or failure of a curricular initiative but be a more complex assessment that provides a fuller understanding of the education process. Parlett and Hamilton (1975) described this as 'illuminative evaluation'.
- The assessment of a curriculum change. A better understanding of a change made to the curriculum can be sought and the extent to which it has been effective measured. The difference between curriculum evaluation and research is sometimes questioned. As argued by Kelly (2004), curriculum evaluation becomes part of curriculum research. In the context of research, the findings of the evaluation must be examined from a generalisable perspective. Curriculum planning and evaluation can be seen as hypotheses to be tested.
- A comparison between and understanding of education programmes in institutions in different countries or with different approaches and admission policies. Such information is useful in the context of student mobility as well as for use in research in medical education.
- The provision of information for potential students. A curriculum evaluation can provide information not only on the excellence of the education programme in a medical school but also on its character, and this might inform a student's choice of medical school. The publication by the Higher Education Funding Councils in the UK of the ratings of the education programmes in medical schools was associated with significant changes in the pattern of schools to which students applied to be admitted.

APPROACHES TO CURRICULUM EVALUATION

A number of frameworks can be used in curriculum evaluation. Here we describe the 'ten questions' framework and the Kirkpatrick four levels of evaluation.

THE TEN QUESTIONS

The ten questions to be addressed in the planning of a curriculum, as described in Chapter 11, can be used also as a framework for a curriculum evaluation. Examples are given of the issues to be addressed relating to each question.

1. *The needs which the medical school aims to meet.* The evaluation of the mission of the medical school is discussed earlier in the chapter.
2. *Learning outcomes.*
 - How is the move to outcome-based education (OBE) interpreted and implemented in the medical school? Teachers have very different perceptions of the significance of learning outcomes and this may influence the delivery of the education programme.
 - A framework to assess the implementation of OBE in a medical school is described in Chapter 9.

3. *The curriculum content.*

- What content is addressed in the curriculum and how does this relate to the stated learning outcomes? How are new subjects such as genetics addressed alongside more traditional subjects?
- What constitutes the core curriculum?
- Is there a danger of information overload and how is this addressed?

4. *Sequence and organisation.*

- What consideration has been given to the organisation and sequencing of content within the curriculum?
- To what extent is there exposure to clinical experiences early in the programme and to the basic medical sciences later?

5. *Educational strategies.*

The 'SPICES' model as described in Chapter 11 is a useful tool to analyse the curriculum and the educational strategies adopted:

- Student-centred/Teacher-centred
- Problem-based/Information-based
- Integrated/Discipline-based
- Community orientated/Hospital orientated
- Electives/Core plus uniform
- Systematic/Opportunistic

Each dimension represents a continuum on which the school will be placed somewhere between the two extremes.

6. *Teaching and learning methods and opportunities.*

- Is there a grid or blueprint that relates the expected learning outcomes to the courses in the curriculum and the available learning opportunities?
- To what extent has there been a move to more 'authentic learning'?
- What is the balance between technology-based and more traditional approaches to learning?
- To what extent is collaborative or peer-to-peer learning encouraged and developed?
- To what extent are the international dimensions to learning incorporated?

7. *Assessment.*

- To what extent is assessment blueprinted against the learning outcomes?
- To what extent is the rich range of tools now available harnessed to assess students' competence?
- Is there a measure of the students' progression through the curriculum?
- Is there adequate feedback to the students and teachers?

8. *Educational environment.*

- How can the education environment in the medical school be characterised and is this in line with the expected learning outcomes?
- Has the education environment been measured using a tool as described in Chapter 18?

9. *Communication about the curriculum.*

- How is the information about the medical curriculum communicated to staff, students and other stakeholders?
- Is a curriculum map available?

10. *Management.*

- Does the management structure support the implementation of the curriculum?
- Are the roles of the teachers defined and matched to their abilities?
- How is teaching recognised and rewarded within the medical school?
- Is a staff development programme implemented and how are teachers kept up-to-date with advances in medical education and in their own specialty?
- What sort of quality assurance processes are in place?

KIRKPATRICK'S FOUR LEVELS OF EVALUATION

Kirkpatrick described four levels for assessing the effectiveness of training. Although the model was developed for use in a business context it has been widely applied in medical education.

Each successive level in the model provides a more precise measure of the effectiveness of the educational programme but requires a more rigorous and time-consuming analysis. Kirkpatrick suggests that evaluation can begin at the first level and, as time and resources allow, move to the higher levels (Fig. 33.1).

Level 1: Opinion/reaction. Evaluation at this level examines how participants in the education programme react to it. Are they satisfied? Do they like or dislike it? Does it meet their needs?

Level 2: Competence/learning. This level moves beyond the learner's satisfaction, and the extent to which there is a difference in the learner's knowledge, skills or attitudes is assessed. Students' performance in written or clinical assessments is studied and taken as evidence of the effectiveness of the education programme.



FIG 33.1 Kirkpatrick's four levels.

Level 3: Performance/behaviour transfer. Changes in the learner's behaviour are studied at this level. As a result of the education programme do students communicate more effectively in their day-to-day contacts with patients? Do students apply their knowledge of the basic sciences to understand better the pathophysiology of patients they encounter?

Level 4: Outcome/results. This level, in the context of Kirkpatrick's work, measures the success of a training programme in terms of the business results. Did sales increase after training was completed? In medical education the question is whether the training of the doctor affects medical practice. Does the training in cardiac auscultation and the interpretation of murmurs, for example, reduce the need for laboratory investigations of patients? Does education of doctors about hypertension reduce the incidence of side effects in patients presenting with raised blood pressure?

UNDERTAKING A CURRICULUM EVALUATION

Undertaking a curriculum evaluation involves:

- Clarifying the purpose of the evaluation. A number of different aims are mentioned earlier in the chapter.
- Deciding the approach to be adopted. The 'ten questions' model and the Kirkpatrick four levels approach are described.
- Engaging the stakeholders. These include the curriculum committee, the teachers, the students, other health professionals, future employers and the public.
- Planning the methods to be used. These almost certainly will include qualitative and quantitative methods.
- Allocating responsibilities for the implementation of the evaluation and the necessary resources to carry it out.
- Planning how the results will be communicated and how action will be taken as a result of the evaluation.

REFLECT AND REACT

1. Are you aware of the stated mission for your medical school or educational body? How successfully is this achieved?
2. What purposes described for curriculum evaluation are relevant to your situation?
3. What steps are taken to evaluate your curriculum and are all the stakeholders involved? Is effective use made of the results of the evaluation?

IF YOU HAVE A FEW HOURS

- Boelen, C., Woollard, B., 2009. Social accountability and accreditation: a new frontier for educational institutions. *Med. Educ.* 43, 887–894.
- Coles, C.R., Grant, J.G., 1985. Curriculum evaluation in medical and healthcare education. *Med. Educ.* 19, 405–422. *A description of what is meant by curriculum evaluation in medical education.*
- Harden, R.M., 1986. Ten questions to ask when planning a course or curriculum. *Med. Educ.* 20, 356–365. *A description of the ten questions to be asked when planning a curriculum.*

IF YOU HAVE MORE TIME

- Berk, R.A., 2006. *Thirteen Strategies to Measure College Teaching. A Consumer's Guide to Rating Scale Construction, Assessment, and Decision Making for Faculty, Administrators, and Clinicians.* Stylus Publishing, Sterling, Virginia. *An amusingly written account of how to assess the effectiveness of teaching.*
- Kelly, A.V., 2004. *The Curriculum: Theory and Practice*, fifth ed. SAGE Publications, London. *A useful account of the purpose, approaches and theory underpinning curriculum evaluation.*
- Kirkpatrick, D., Kirkpatrick, J., 2006. *Evaluating Training Programs: The Four Levels*, third ed. Berrett-Koehler, San Francisco. *A description of the Kirkpatrick four levels of evaluation.*
- Parlett, M., Hamilton, D., 1975. *Evaluation as illumination.* In: Tawney, D. (Ed.), *Curriculum Evaluation Today: Trends and Implications.* Schools Council Research Studies, Macmillan, London. *A classic description of illuminative education.*