

# 4

## Assessing learning

*Timothy Rennie*

Assessment is an established and often ongoing part of most of our lives. Assessment goes hand in hand with education, training and learning. Therefore, most people have been assessed at some point, and usually from an early age. As a teacher you will inevitably be presented with an opportunity to mark, grade or indicate the level of competence or performance of your students through assessment. The purpose of this chapter is to give you a quick guide to the core principles of assessment and to understand the why, what and how. It also presents a selection of assessment techniques that can be used individually or together in different settings.

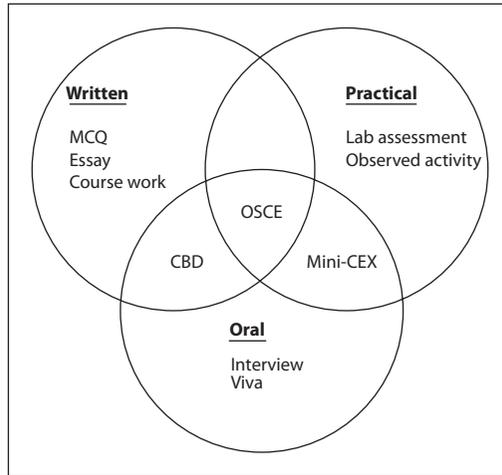
### **What is assessment?**

Assessment is often thought of as a way to quantify or qualify previous learning and/or direct future learning. Another way to help define the concept of assessment is by providing examples of how it is conducted.

Methods of assessment can be classified in a number of different ways. In Fig. 4.1 they are categorised as oral, written or practical. You can see from the diagram that there can be a cross-over between these various methods of assessment. For example, the **Objective Structured Clinical Examination** (OSCE – see the appropriate section later in this chapter for more information) usually incorporates oral and written elements as well as practical skills.

Assessment methods may also be described as quantitative or qualitative in nature. For example, a **Multiple Choice Question** (MCQ) is a pre-defined, sometimes pre-validated method and will have pre-defined answers. So structured is the nature of MCQ assessment that they are sometimes marked or graded by automated machines.

Conversely, an oral viva voce examination (meaning ‘by word of mouth’) will usually be based on the student’s dissertation or thesis. While there are guidelines for conducting viva examinations, the examiners will have a large degree of subjective freedom – and necessarily so – to question, probe, discuss and debate. This may relate to information not included in the thesis, such as the decisions made by the learner, and even their



**Figure 4.1.** Methods of assessment.

motivations or aspirations in conducting their work. The examiners will have no direct control over the responses that the learner delivers and, therefore, the answers cannot be pre-defined.

### **Formative, summative and ongoing assessment**

Most learners experience some degree of **ongoing assessment**, be it a test at the beginning of the class on the previous lesson or structured formal modular examinations throughout a course or degree. The marks may or may not contribute to a final mark – they may just indicate how well you are progressing. They may be used to deliberately motivate learning, or this could just be an unintended side-effect. However, the concept of ongoing assessment does not really capture this motivational element, whether the assessments are intended to inform future learning or simply provide an indication or ‘benchmark’ for how much learning has happened.

Another way of defining types of assessment is whether they inform learning progression or whether they provide an indication of how much learning has taken place. **Formative assessment** *informs* learning, whereas **summative assessment** is the *sum* of learning.

From this perspective a single assessment method may be used for different purposes. It may be used to indicate how learners are progressing throughout a learning experience (formative) and in this way point to strengths or gaps in knowledge or skills. This in turn can direct the focus of the learner’s learning. Conversely, the assessment may be used as an end-point (summative) examination, such as an end-of-year assessment in a degree course. However, in both these examples the assessment can be described as ‘ongoing’ despite serving different purposes.

There are some examples of assessment tools at the end of this chapter that can provide a comprehensive assessment of an individual's learning over a period of time (see 'Common assessment methods used in health professions education'). To help you decide which method of assessment is most suitable for your learners it is useful to ask yourself these questions:

- Why am I assessing, what is the purpose?
- What am I assessing against?
- Who is being assessed and who is the assessor?
- At what point in the learner's education or development is the assessment taking place?
- In what setting will the assessment be performed?
- How do I judge what is 'good'?
- How do I feed back the results to the learner?

## Why assess?

Questioning why we are assessing may lead us to consider the philosophy and ethics of assessment, which we will not go into here. The purpose of this chapter is to outline the pragmatic perspective: understanding the 'why' may inform better the 'how' and the 'what' (see Exercise 4.1).

Assessment is used to demonstrate a level of **competence** (see Box 4.1). This can be particularly important if someone's competence has a profound and direct impact on others (for example, in medical practice). Assessment may also be part of a selection process such as degree entry examinations or interviews for employment. Assessment is also used as a management tool – to provide evidence of incompetence or poor practice, or to demonstrate value and stimulate activity or competition.

There may also be other, more ignoble, motivations for assessment. Assessment could (but should not) be used to catch people out or to deliberately set levels that people cannot attain in order to exclude or demote. This highlights the importance of assessment being valid, practical and fair.

## What am I assessing against?

You may already be able to answer the question why you are assessing. The overall aim may be to demonstrate competence or simply to show that learning has occurred. However, it may not be immediately clear what you are assessing against. There may be a general **curriculum** – this will usually give a broad summary of what is expected of learners and, therefore, what should be assessed. Outside the academic setting you are more likely to find competencies – whether defined or not – that are expected to be attained for an individual to carry out their job or role sufficiently.

*Box 4.1 Some definitions (UK setting)*

*Competencies:* ‘Competencies are descriptors of the performance criteria, knowledge and understanding that are required to undertake work activities. They describe what individuals need to do, and to know, to carry out a particular job role or function, regardless of who performs it.’

*(Competence Framework, NHS Connecting for Health)*

*Assessment:* ‘The process of measuring an individual’s progress and accomplishments against defined standards and criteria, which often includes an attempt at measurement. The purpose of assessment in an educational context is to make a judgement about mastery of skills or knowledge; to measure improvement over time; to arrive at some definitions of strengths and weaknesses; to rank people for selection or exclusion, or perhaps to motivate them.’

*(Postgraduate Medical Education and Training Board)*

*Evaluation:* ‘Evaluation is the collection of, analysis and interpretation of information about any aspect of a programme of education or training as part of a recognised process of judging its effectiveness, its efficiency and any other outcomes it may have.’

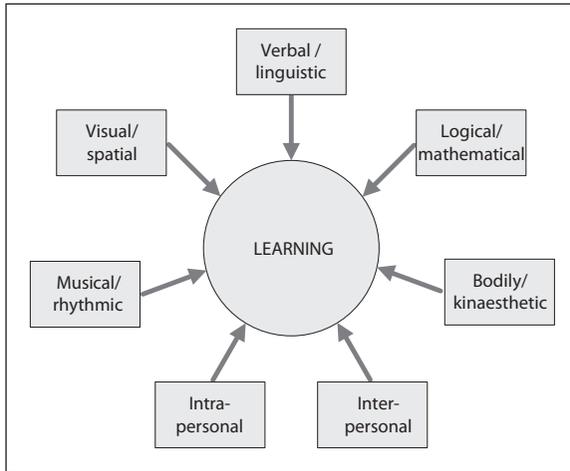
*(Thorpe 1993)*

Both curricula and competencies may be articulated into specific **learning outcomes** (what is expected to be learned) and/or **learning objectives** (what and how learning takes place) (see Chapter 2). A short course, for example, is likely to have an overall aim to improve competence or educate in a particular area and some specific objectives that indicate what learning should take place.

Consider a two-hour first aid course. The overall aim will be to help the learner to be better at, or more informed about, delivering first aid. Within this there may be specific learning outcomes such as:

- 1** Be able to identify the appropriateness of cardiopulmonary resuscitation (CPR) for a person who is unconscious
- 2** Be able to deliver CPR when appropriate.

In general, assessment should relate to specific indicators of what learning or level of competence is expected. In the absence of these, learning outcomes can be derived from the overall aims or curriculum of the learning programme.



**Figure 4.2.** Multiple intelligences. (Adapted from Gardner 1983.)

In some instances the aims, learning outcomes and learning objectives may be defined by the learner (consider again the PhD thesis and viva). However, if the learning aims or objectives are not defined when they are expected to be, are defined poorly, or do not relate well to the work undertaken, this will very likely be a criticism by examiners.

The examples of assessment detailed in this chapter focus on cognitive ability, but it is worth remembering other aspects of performance that may be assessed, such as physical ability or interpersonal skills (see Fig. 4.2). Taking again the example of the first aid course, assessment could demonstrate cognitive recall of information as well as competence in enacting a typical first aid scenario. However, concepts that are demonstrated physically may also be assessed theoretically – a chemical titration that exhibits a colour change can be described regardless of whether it is actually conducted. Indeed, assessment may be used to quantify or qualify learning that is difficult to demonstrate but can be described instead.

## How assessment meets learning

In an ideal world the best assessment would be one that most fairly reflects the learning, competence or performance of an individual. However, apart from the practical constraints of time and personnel, we must also appreciate that no method of assessment will assess the entirety of what has been, or is being, learned. People exhibit different characteristics and intelligences – be they learned or innate – and develop different **learning styles** and strategies. (See Chapter 2 for more information on learning styles.)

Some people may rely on memory and recall whereas others may rely on conceptual thinking or ‘stories’. Any given assessment will have a bias in some way towards a particular learning style or even human characteristic. Therefore, rather than trying to suit the ‘horse to the course’ (which could be criticised as being unfair in itself and impossible anyway) it is worth at least trying to create a level playing field in delivering assessments.

One way of doing this is to create a mix of different assessment methods that measure learning from different aspects. This can also give the assessor greater confidence in the results, provided that the assessment methods are robust and are truly measuring what is intended. Sometimes only one assessment method will be suitable to measure a very specific element of learning. Similarly, there will be some aspects of learning that cannot be assessed accurately or reliably. Although we should strive to be creative in our methods, this has to be an accepted limitation.

## The assessor

The role of the assessor is very important. There ought to be objectivity in relation to the learner, whether they come face-to-face or not. There should be no conflict of interest, in order to eliminate positive or negative bias. This can be tricky as in some circumstances the ideal person to assess is the person who is closest to the learner’s practice or performance. As an assessor you should ask yourself, do you have multiple interests in the outcome of the assessment that may impact on its outcome? Discussion with colleagues or the human resources team could help to answer this.

The roles and responsibilities of the assessor are varied. Again, in the absence of guidelines or a pre-defined set of answers, it is helpful to consider the purpose of the assessment and focus on the learning objectives to clarify what should be assessed. If the assessor plays an active part in the interpretation of the learner’s responses then some degree of competence on the assessor’s part and/or familiarity with both the assessment method and responses may be expected. Indeed, specialist assessors may be specifically recruited to perform or mark the assessment.

## Assessment value

How do you know whether your judgement of a learner’s performance is accurate? Of course there is often subjectivity in assessment and assessors must ensure that outcomes can be reproduced to a standard degree of error (that is, different assessors will give similar marks). Even if the assessment is a **subjective assessment** there may be some structured criteria and recording format that you are expected to follow in order to provide feedback, either for the learner or for quality assurance purposes. If the outcome of the assessment is a quantitative mark or percentage, the question arises of what

**Table 4.1** Objective and subjective rating scores

Criterion	Highly competent		Competent		Not yet competent/ does not meet minimum criteria		Unacceptable	
	<b>Overall ability</b>	Evidence presented <i>substantially</i> exceeds minimum requirements of competence in terms of content, structure, knowledge, insight and other defined criteria		Evidence presented is a sufficient, valid, reliable indication that the learner has met minimum requirements of competence		Evidence presented indicates that the learner has not yet demonstrated overall competence or is not yet competent in terms of one or more of the specified criteria		Evidence presented is substantially flawed, gross inaccuracies, omissions, irrelevant content or substantial deviation from the specified criteria
	Excellent	Very good	Good; Competent	Satisfactory	Just below standard	Weak	Very weak	Very little of relevance
Mark earned	85–100%	75%	65%	55%	45%	35%	25%	0–15%

this actually means or equates to. This may be easier to judge if there is a pass mark or if percentage points are categorised into different levels of proficiency.

An example of objective and subjective scoring is shown in Table 4.1. There is subjective categorisation and there are equivalent descriptions of levels of competence and percentage scores.

It is also worth remembering that the value of the outcome may only be valid for a period of time. The significance of any achievement is likely to decrease over time or through greater achievements. For example, it is likely that qualifications achieved at school are superseded by university qualifications, and they in turn by professional qualifications – although basic, lower-level qualifications may still be required to demonstrate proficiency in core subjects such as English and mathematics. If there is a specific expiry date, it may relate to ensuring that individuals are up-to-date in their learning or competence – consider again the first aid course.

So how can we ensure that assessments are a true reflection of performance? In the United Kingdom there is annual debate following the release of key school examination results concerning their validity. As results continue to improve year on year, the media perception is that the examinations are getting easier. In certain circumstances, particularly where one is being measured against one's own previous performance in formative assessments, this is irrelevant – you would expect and hope to see improvement over time.

However, there may be justification for reviewing a cohort of results by analysing, for example, the ‘average’ score (mean/median/mode), the standard deviation from the mean, and any outlying results that may have a disproportionate effect on the average. How does this year’s analysis compare with those of the previous year(s)? Are there any patterns? If so, can these patterns be explained? How will this inform how we assess next year?

If the assessment method included a battery of questions or ‘items’, such as the MCQ, a technique called a ‘test-item analysis’ may be employed, particularly in an academic setting. This judges each item against pre-defined external criteria or against the remaining items in the assessment. In performing this analysis, there may be one item that is consistently answered incorrectly or not answered at all, even by the highest scorers. This raises suspicion as to whether the question was fair or whether it could be reasonably interpreted in a number of ways.

Two outcomes may result from the analysis. One is the *post-hoc* improvement in the assessment questions or items. The second is adjustment to the assessment scores to reflect the difficulty or fairness of the assessment. If the second option is used, it should be done consistently (i.e. with every assessment of its kind) and in a disciplined way against agreed criteria in order to avoid accusations of manipulating results.

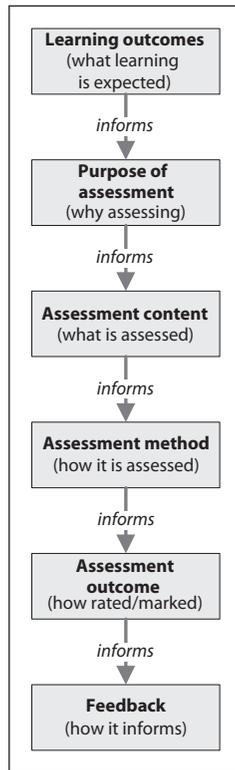
## Feedback on assessment

The question of how the outcome of assessment is fed back to the learner again relates to the nature of the assessment. If the assessment is formative, it is likely that learners will receive feedback in order to guide future learning (see Fig. 4.3). However, if the outcome of the assessment is summative, feedback may not be given – often not even under appeal – especially if the desired outcome is achieved. This usually relates to the practical balance of providing feedback versus the transparency of the academic process. Academics need to be transparent, but the reality of providing hundreds or thousands of individuals with detailed feedback is not viable. However, educational institutions will typically have processes in place for appeal and re-marking even if feedback is not granted.

There is also the danger of conflict if assessment is heavily subjective and the learner disagrees with the assessor’s interpretation; this is not easily managed on a large scale.

## Setting

The setting or venue for an assessment is also important. The setting may not be decided until the day of the actual assessment, for example if it is a work-based observation. In this instance the assessor should always take



**Figure 4.3.** The process of assessment.

the lead in ensuring that the setting is appropriate. Is there enough space? Is it too busy or noisy to properly conduct an assessment? If patients or members of the public are involved, are there any confidentiality issues? Is it possible to provide feedback within this setting if it is required?

## Plagiarism in assessment

It is extremely important to recognise the significance of plagiarism in assessment as the risk of accrediting learning that has not taken place could result in tangible damage even to human life. Individual institutions will provide their own guidance relating to plagiarism in assessment and we refer the reader to these institutional policies and practices. The process for tackling plagiarism should nonetheless be consistent, fair and transparent. However, the message to be encouraged regarding plagiarism is that it is never in the learner's interests, regardless of whether it is detected or not. In particular, with formative assessment – where the only direct benefit is to the learner – plagiarism defeats the purpose of this method and therefore actually disadvantages the learner.

It should also be recognised that plagiarism is not always intentional. For example, learners may not be aware that they must reference texts or quotations that they have used. There is a need for clear understanding at the commencement of learning as to what plagiarism encompasses. At this stage it is also advisable to encourage an open culture of reporting plagiarism to ease detection. There are a number of detection methods and any one method should not be relied upon. With the increasingly broad range of assessment methodologies there is a need for detection methods to respond. For further information on plagiarism see the good practice guide by Carroll and Appleton (2001).

## Methods of assessment

A vast battery of assessment methods have emerged that seek to provide a means of quantifying and articulating learning. The following methods have been selected to demonstrate the broad field of tools that can be applied in different learning settings and for different reasons; for example, some that are competency-based and some that are not. The list does not seek to be exhaustive – further reading suggestions can be found at the end of the chapter.

### Essay

Essays are a common method of formal assessment, especially in the academic field. Written from the perspective of the author, they usually offer critique or argument around a central theme. Examples of types of essay include short-answer essays, standard essays, extended-time essays, and dissertations (see Miller *et al.* 1998). For the purposes of assessment this theme may be posed as a question or statement to which the learner responds. When setting an essay as an assessment you may like to consider the following:

- Is an essay the most appropriate form of assessment to use? Consider the skills and resources that are required for essay writing and assessment.
- How important is it that the essay is written in an invigilated setting? Consider the risk of plagiarism and the extent to which sources or evidence need to be referred to.
- Is the essay being used as a summative or formative assessment? This may inform the degree of feedback that is required or desired.

There is a great degree of flexibility and subjectivity that can be daunting both for the learner and the assessor. However, in broad terms, there are a number of criteria against which an essay can be judged. When marking an essay, consider the following points: