Multiprofessional Faculty Development - Workplace Based Assessment
What is workplace-based assessment?

Workplace-based assessment refers to the assessment of working practices based on what doctors actually do in the workplace, and is predominantly carried out in the workplace itself (PMETB, 2007).

Miller’s ‘pyramid of competence’ (Miller, 1990) is also useful for mapping assessment methods against the various tiers of the pyramid. The ‘knows’ level of the pyramid can be assessed using simple knowledge tests, e.g. multiple-choice questions (MCQs). The ‘knows how’ level can be assessed using unfolding patient management problems (PMPs) or essay questions. Objective Structured Clinical Examinations (OSCEs) can assess the ‘shows how’ level. The difficulty has always been assessing the ‘does’ level, which in professional practice refers to performance in context.

Miller’s pyramid

Competence versus performance

Competence indicates what people can do in a contextual vacuum, under perfect conditions. This might be evident using controlled assessment methods looking at the lower tiers of Miller’s pyramid. Performance, however, indicates how people behave in real life, on a day-to-day basis.

The problem is that what doctors do in controlled assessment situations correlates poorly with their actual performance in professional practice (Rethans et al., 2002). Therefore we need assessment methods that focus on the top end of the pyramid. This is where workplace-based assessment comes in.

In workplace-based assessment we are assessing performance in vivo using samples of data gathered from the working practice of the doctor. The notion of adequate sampling is critically important and will be referred to later.

Thinking point

- Are you using, or familiar with, any assessment methods that assess at the top level of Miller’s
pyramid?
Principles of assessment

The traditional educational paradigm below indicates the importance of integrating teaching, learning and assessment, where assessment should figure as an integral part of educational planning. However, it is all too easy to view assessment as a ‘bolt-on’ component at the end of an educational programme.

Utility

The utility, or usefulness, of an assessment has been defined as a product of its reliability, validity, cost-effectiveness, acceptability and educational impact (van der Vleuten, 1996). In later years the term ‘feasibility’ has been added. Utility can be expressed as an index:

\[
\text{Utility} = \text{educational impact} \times \text{reliability} \times \text{validity} \times \text{cost effectiveness} \times \text{acceptability}
\]

Utility can be applied to an entire assessment system or to an individual assessment method or component of the system. The concept is important in that no single element of utility should be regarded as a panacea. Instead, assessment design should pay attention to all the elements within the utility equation, although it is recognised that there may be a ‘trade off’ between the elements.
Thus, traditional approaches to maximise the reliability or reproducibility of assessments can have a negative educational impact on the learner by reducing the opportunity for meaningful developmental feedback. Whereas workplace-based assessments can offer high educational impact but might not be performed as reliably as other medical assessments, e.g. MCQs.

**From methods to programmes**

Traditional approaches to developing medical assessments have been founded on the notion that domains of medical competence (e.g. problem solving, communication skills) are stable and generic. As a result, it is possible to design tests that assess these domains separately and reliably. This has led to the ‘one trait, one instrument’ approach to assessment (Schuwirth and Van der Vleuten, 2004).

However, there is a growing realisation that medical competence is specific to particular clinical situations or clinical contexts, and not generic. This is called case specificity and poses particular problems in assessing doctors. To overcome this, the most important thing is to sample widely across the content of the curriculum.

In recognition of the complexity of assessing professional competence there is a growing realisation that assessment should be considered as a programme of activity requiring the assimilation of quantitative and qualitative information from different sources. Assessing doctors in their actual working environment therefore offers the opportunity to gather sufficient information to build up a ‘rich picture’ of them.

This also has a powerful impact on work-based learning through the provision of feedback and encouragement of reflection, and takes us right back to the educational paradigm.

**Thinking point**

- Try applying the utility index to an assessment method with which you are familiar. What are the assessment’s strengths and weaknesses?
Why workplace-based assessment?

The educational argument for integrating teaching, learning and assessment is powerful. We know that assessment drives learning and it is therefore imperative that workplace-based assessment focuses on important attributes rather than what is easiest to assess. Complex professional attributes are difficult to assess using standardised assessment methods such as written exams or OSCEs. These are better assessed in workplace situations.

We know that learning is at its most powerful when it is ‘authentic’. This means when the learning is integrally related to the understanding and solution of real-life problems (Kaufmann et al., 2000). We also know that assessment is more valid the closer we get to what we wish to assess. Authenticity is particularly important when it comes to assessing doctors because of the problem of case specificity. Therefore observing doctors in real-life situations becomes more important, allowing the collation of multiple snap shots of performance to give a global and holistic perspective of the learner.

The Postgraduate Medical Education and Training Body (PMETB) has the responsibility for overseeing specialist training and recognises the importance of reassuring the public about the safety and competence of its doctors. It has recommended in its principles of assessment that workplace-based assessments form an important part of the assessment framework (PMETB, 2003).

Workplace-based assessments cannot replace standardised assessments yet. There are problems with reliability as a result of inconsistent application of tools across different raters or assessors. There is potential conflict in the role of an educational supervisor who is supervising the learner, but is also involved in the assessment process. There are also problems of attribution to an individual, as opposed to the clinical team, when routinely collected clinical practice data are assessed.

So what issues are important to consider in designing a programme of workplace-based assessment?
Deciding what is to be assessed

The areas chosen to assess in workplace-based assessment are usually articulated as a series of competencies. These should be blueprinted against the curriculum, and most importantly encourage trainee development. Let us look at those three issues in a little more detail.

Competency based

Workplace-based assessment is usually competency based. Despite recent criticisms of competency-based education as a whole (Talbot, 2004), concerns are usually expressed where competencies are viewed as narrow, reductionist and overly simplistic. Competencies used for the purpose of designing workplace-based assessments are best written as holistic statements which are framed as ‘a complex structuring of attributes needed for intelligent performance in specific situations’ (Gonczi, 1994).

The following 12 competencies serve as an example and have been developed for use in the workplace-based assessment component of the nMRCGP, the exit assessment for general practice training.

<table>
<thead>
<tr>
<th>Workplace-based assessment competencies for the nMRCGP (RCGP, 2007)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Communication and consultation skills: this competency is about communication with patients, and the use of recognised consultation techniques.</td>
</tr>
<tr>
<td>2. Practising holistically: this competency is about the ability of the doctor to operate in physical, psychological, socioeconomic and cultural dimensions.</td>
</tr>
<tr>
<td>3. Data gathering and interpretation: this competency is about the gathering and use of data for clinical judgement, the choice of physical examination and investigations, and their interpretation.</td>
</tr>
<tr>
<td>4. Making a diagnosis/making decisions: this competency is about a conscious, structured approach to decision making.</td>
</tr>
<tr>
<td>5. Clinical management: this competency is about the recognition and management of common medical conditions in primary care.</td>
</tr>
<tr>
<td>6. Managing medical complexity and promoting health: this competency is about aspects of care beyond managing straightforward problems, including the management of co-morbidity, uncertainty, risk and the approach to health rather than just illness.</td>
</tr>
<tr>
<td>7. Primary care administration and IMT: this competency is about the appropriate use of primary care administration systems, effective record keeping and information technology for the benefit of patient care.</td>
</tr>
<tr>
<td>8. Working with colleagues and in teams: this competency is about working effectively with other professionals to ensure patient care, including the sharing of information with colleagues.</td>
</tr>
<tr>
<td>9. Community orientation: this competency is about the management of the health and social care of the practice population and local community.</td>
</tr>
<tr>
<td>10. Maintaining performance, learning and teaching: this competency is about maintaining the performance and effective continuing professional development of oneself and others.</td>
</tr>
<tr>
<td>11. Maintaining an ethical approach to practice: this competency is about practising ethically with integrity and a respect for diversity.</td>
</tr>
<tr>
<td>12. Fitness to practise: this competency is about the doctor’s awareness of when their own performance, conduct or health, or that of others, might put patients at risk and the action taken to protect patients.</td>
</tr>
</tbody>
</table>
Blueprinted in the curriculum

To ensure that assessments are integrated with the curriculum, competencies chosen for assessment should map directly back to the curriculum to ensure that the curriculum is being adequately covered, and that the assessment methods used sample across all the competencies and therefore there is widespread sampling of the curriculum. See attached Teachers’ Toolbox item on Assessment blueprinting.

Developmental

As already discussed, workplace-based assessment offers the opportunity to link teaching, learning and assessment effectively, and the developmental nature of the assessment should therefore be a key feature.

To enhance educational impact, the use of holistic competencies and a developmental continuum is recommended. Developmental progressions in the literature, such as that described by Dreyfus and Dreyfus (1986) may be helpful in constructing the developmental continuum. These progressions work through the following stages to from novice to expertise:

- novice
- advanced beginner
- competent
- proficient
- expert.

The figure below gives an example of how progression statements can be developed in relation to a competency. In this case the competency is that of ‘practising holistically’ taken from one of the 12 competencies of workplace-based assessment for the nMRCGP assessment of the RCGP.

Example of a competency progression statement (RCCP, 2007)

<table>
<thead>
<tr>
<th>2</th>
<th>Practising holistically</th>
</tr>
</thead>
<tbody>
<tr>
<td>This competency is about the ability of the doctor to operate in physical, psychological, socioeconomic and cultural dimensions, taking into account feelings as well as thoughts.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Insufficient evidence</th>
<th>Needs further development</th>
<th>Competent</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>From the available evidence, the doctor’s performance cannot be placed on a higher point of this developmental scale.</td>
<td>Enquires into both physical and psychological aspects of the patient’s problem.</td>
<td>Demonstrates understanding of the patient in relation to their socioeconomic and cultural background.</td>
<td>Uses this understanding to inform discussion and to generate practical suggestions for patient management.</td>
</tr>
<tr>
<td>Recognises the impact of the problem on the patient.</td>
<td>Additionally, recognises the impact of the problem on the patient’s family/carers.</td>
<td>Recognises and shows understanding of the limits of the doctor’s ability to intervene in the holistic care of the</td>
<td></td>
</tr>
</tbody>
</table>
Uses themself as the sole means of supporting the patient.

Utilises appropriate support agencies (including primary healthcare team members) targeted to the needs of the patient.

Organises appropriate support for the patient’s family and carers.

Such a continuum has the advantage of explicitly illustrating the direction of travel for trainees, rather than merely pointing out the level below which they should not fall.

This supports the concept of ongoing evidence collection throughout the training period, but with regular, well-circumscribed staging reviews at which the developmental framework is reviewed and the learner’s progress through it judged.

So workplace-based assessment must provide detailed formative and developmental feedback to the learner. This raises the tension of potentially mixing formative and summative elements, but it is possible to address this through the careful design of the assessment system. Separating the interpretation of evidence from its elicitation is one way around the problem (William and Black, 1996).

**Thinking point**

- In practical terms what might you do to address the tension between the formative supervisory role and the summative assessment requirements of being an educational supervisor?
Deciding how to assess

Collecting the evidence

Collecting ‘sufficient’ evidence is essential in making a judgement about the attainment of competence. Sampling widely across a number of clinical and contextual situations is important to overcome the problem of case specificity.

In the assessment of ‘work’, in contrast to traditional methods there is no single ‘controlled’ method that can be developed. A variety of sources of information will be needed, which gives rise to the notion of a ‘tool-box’ of approved methods (see below). In considering the individual tools it is worth recognising that, even unstandardised, they can be made sufficiently reliable, provided enough sampling occurs, and the tools are used sensibly and expertly (van der Vleuten and Schuwirth, 2005). However, it is important to remember that the tools form part of the overall assessment programme. Attention should focus on the reliability of the entire programme of assessment, not just the individual tools themselves.

From a feasibility point of view, it is likely that a lot of workplace-based assessment activity will occur locally, i.e. within the workplace setting. This calls for effective educational supervision. An educational supervisor should be responsible for co-ordinating the assessments, but the process of collecting the evidence should be learner led.

The educational supervisor can also be involved in making judgements about the learner’s progress against the competencies provided. The potential role conflict for the educational supervisor responsible for facilitating progress and making judgements on it can be overcome, provided the differing roles are made explicit to the learner. In other words, when it’s assessment time, the learner needs to know, and be adequately prepared for it.

Triangulation

The confidence in the reproducibility of judgements in a workplace assessment can be improved through triangulation within the workplace assessment. This involves using different methods to collect evidence and using the views of multiple raters over a period of time.

Triangulation with other assessments external to the workplace is also important. An overarching assessment strategy is essential, in which workplace assessment is supported by rigorous tests, e.g. those of ‘knowledge’ and ‘skills for clinical method’. It is recommended that there should be an overarching assessment strategy (i.e. a programme for assessment) for the whole training period, and that this is blueprinted to the curriculum. Workplace assessment and other appropriate methods form part of this.

Thinking point

- Think of a professional competency area from your own specialty, or the Foundation curriculum. What do you think would constitute ‘sufficient evidence’?
Workplace-based assessment and the portfolio

A portfolio can be seen as a dossier of evidence collected over time, which demonstrates a doctor’s education and practice achievements (Wilkinson et al., 2002).

In essence, a portfolio, if well constructed, should describe the journey of a learner towards the attainment of professional competence.

The box below demonstrates the key functions of a portfolio and how it relates to workplace-based assessment.

Functions of a portfolio

- It aims to serve as the reflective learning log of the learner, available to be shared with their educational supervisor.
- It demonstrates the learner’s progress towards covering the breadth and depth of the curriculum.
- It acts as a repository for assessments carried out as part of workplace-based assessment.
- It acts as a framework for the learning agreements between learners and teachers.
- It charts a learner’s progression and ‘learning journey’, and can help in making career choices and decisions.
Some methods for gathering evidence

A brief overview of some of the main methods for gathering evidence currently used in the United Kingdom is shown below.

Observations of patient encounters

Mini-CEX (the clinical encounter)

It has been traditional for clinical skills to be assessed by the ‘long case presentation’. The problem of case specificity using this technique, limiting the potential to sample widely across different clinical cases and contexts, has given rise to a version of the ‘short case examination’, popularly known as the mini-CEX (Norcini et al., 1995). The technique has been developed to assess the clinical skills that trainees most often use in real clinical encounters. It is based on assessment of multiple encounters within a hospital setting observed by an educational supervisor or other clinician.

Direct observation of skills

This category is again an assessment of real-life activities where the focus of the assessment is the skill with which the activity was performed, e.g. direct observation of technical skills (DOPS), teaching skills and presentation skills. ‘The consistent feature is that one or more assessors, who are trained in the assessment of that skill, make a judgment about a real life performance’ (PMETB, 2007).

Discussion of clinical cases

Case-based discussion (CbD)

The technical development for the use of CbD in UK training assessment systems is based on the use of case-based discussion in the General Medical Council (Southgate et al., 2001), which itself is derived from chart-stimulated recall oral assessments used extensively in the US and Canada. CbD is one of the evidence gathering tools used in the framework for workplace-based assessment in the UK Foundation programme and is also being used in specialty training programmes, e.g. the nMRCGP licensing assessment of the Royal College of General Practitioners.

Performance data


- The basis for making a judgement:
  - outcomes of care, while being the most desirable, are limited by problems of attribution (to the individual), complexity, case mix and numbers
  - process of care is more directly in the control of the doctor but does not necessarily guarantee the best patient outcome
  - volume of activity information is premised on the basis that the more of a given activity that a doctor performs, the better the likely quality of care in relation to that activity. This method of judgement is more likely to be applicable to surgeons who are engaged in numbers of specific technical procedures.

- The actual methods of collecting data:
  - external audit of medical records
  - use of administrative databases
Multi-source feedback

Lockyer (2003) describes the principles involved in multi-source feedback (MSF), which is seen as a practical approach to assessing doctors in the workplace. The goal is to view a person’s work from a variety of perspectives. In medical settings, physician colleagues (peers), co-workers and patients can be asked to complete surveys about the doctor.

The person being assessed receives feedback with their own aggregate ratings, along with ratings for others being assessed at the same time. The opportunity for comparing self-assessment data with those provided by raters is clear. Early evaluation suggests that 6–8 raters per cycle of MSF are required to make the tools sufficiently reliable.

MSF tools can be further classified into:

- peer–rating tools, e.g. the TAB (Team Assessment of Behaviour) which is used in foundation training
- patient satisfaction questionnaires (PSQs). The following list indicates some of the tools being used within the UK at present (Chisolm and Askham, 2006):
  - Consultation and Relational Empathy (CARE) questionnaire
  - SHEFFPAT (Sheffield patient assessment tool)
  - General Practice Assessment Questionnaire (GPAQ)
  - Consultation Satisfaction Questionnaire (CSQ)
  - Doctors Interpersonal Skills Questionnaire (DISQ).

Thinking point

- Think of a professional competency area from your own specialty, or the Foundation curriculum. What do you think would constitute ‘sufficient evidence’?
Quality assurance of a programme of workplace-based assessment

In viewing the quality assurance of a programme of workplace-based assessment it is helpful to review the utility equation described earlier. Workplace-based assessment has strengths in the areas of validity (by virtue of its authenticity), educational impact and acceptability (because it reconnects teaching and learning), and feasibility (through local assessment).

There are, however, problems with demonstrating its reliability using traditional psychometric approaches. As Southgate et al. (2001) point out, ‘establishing the reliability of assessments of performance in the workplace is difficult because they rely on expert judgements of unstandardized material’.

In workplace assessment like any other form of assessment there are several potential threats to reliability:

- inter-observer variation (the tendency for one observer to mark consistently higher or lower than another)
- intra-observer variation (the variation in an observer’s performance for no apparent reason - the ‘good/bad day’ phenomenon)
- case specificity (the variation in the candidate’s performance from one challenge to another, even when they seem to test the same attribute) (Crossley et al., 2002).

Therefore it might be helpful to reframe the term reliability with the phrase maximising ‘consistency and comparability’. Baker et al. (1991) have proposed a framework that can do this by incorporating the following five elements within the assessment design.

- **Specification** of standards, criteria, scoring guides.
- **Calibration** of assessors and moderators.
- **Moderation** of results, particularly those on the borderline.
- **Training** of assessors, with retraining where necessary.
- **Verification and audit** through quality assurance measures and the collection of reliability data.

It is clear, therefore, that the implementation of workplace-based assessment will require a complementary training programme, arrangements for calibration, a procedure for the moderation of results and a raft of quality control and reliability checks.

The more that teachers can be engaged in assessment, in selecting methodologies, generating standards, discussing criteria, etc., the more they will be empowered in the educational process.

**Thinking point**

- What do you think are the training requirements for trainers and supervisors in relation to the above tools?
To sum up

Workplace-based assessment offers the opportunity to connect teaching, learning and assessment, and can assess the ‘does’ dimension of Miller’s pyramid of competence. It is a useful component of an overall assessment programme.

When designing a programme of workplace-based assessment it is important to address the following issues.

- The utility of the entire assessment programme.
- Clarity on what is being assessed is needed through the identification of holistically described professional competencies.
- Attention should be given to the developmental nature of the assessment and this should be built into progression statements for each competency.
- A variety of assessment tools should be used to gather evidence from multiple clinical contexts using multiple raters.
- All the evidence should be collated, synthesised and judged at regular intervals by an educational supervisor to assess the learner’s progress.
- The consistency and comparability of workplace-based assessments should be maximised through a programme of quality assurance.

Congratulations

You have now reached the end of the module. Provided you have entered something into your log you can now print your certificate. To generate your certificate please go to ‘my area’ and click on ‘complete’ in the course status column. Please note, you will not be able to print your certificate unless you have entered something in your ‘reflections area’.

Please now take a moment to evaluate the course and enter your comments below.
Further Information

This module was written by Dr Nav Chana, senior lecturer at the Faculty of Medicine and Biomedical Sciences St Georges University of London and Associate Director for Postgraduate General Practice Education London Deanery. The module relates to areas 1, 3 and 4 of the Professional Development Framework for Supervisors in the London Deanery.

References


Course Glossary

Aim
An aim in educational terms, is a brief statement of intent, indicating the scope and range of intended learning outcomes that the educational episode has been structured to address.

Assessment
Assessment is the term used to indicate an appraisal of students’ performance. Typical formal assessments in medicine include written examinations, Multiple choice questionnaires (MCQ), observations of clinical or communication skills, Objective Structured Clinical Examinations (OSCEs) and Multi-Source Feedback (MSF). Assessments may be summative (where the marks gained contribute to a formal grade or award) or formative (where the focus is on providing feedback for ongoing development).

Class
Class refers to hierarchical differences between individuals or groups in societies or cultures. Factors that determine class may vary widely from one society to another. However, economic disadvantage and barriers to access services are major issues within class discrimination.

Competencies
In assessment terms competencies refer to a set of professional abilities that includes elements of knowledge, skill, attitudes and experience. Competencies are similar to objectives and outcomes in that they provide a means of specifying attributes in relation to the ultimate intended performance that the competencies underpin (Grant, 2007, p 21). The use of competencies has been widespread in practical vocational subjects such as healthcare, management and engineering. Competence based curricula can be used as a basis for learning and teaching, for assessment and to help ensure professional accountability. Programmes for professions such as medicine usually include specific practical competences and the integration of more complex skills, knowledge and behaviours.

Curriculum
A detailed schedule of the teaching and learning opportunities that will be provided (GMC, 2004)

Curriculum
The GMC, 2004 described the curriculum as a detailed schedule of the teaching and learning opportunities that will be provided. A curriculum is a statement of the aims and intended learning outcomes of an educational programme. It states the rationale, content, organization, processes and methods of teaching, learning, assessment, supervision, and feedback.

Diversity
Diversity literally means difference. When it is used as a contrast or addition to equality, it is about recognising individual as well as group differences, treating people as individuals, and placing positive value on diversity in the community and in the workforce. Historically, employers and services have ignored certain differences.

Feasibility
This identifies the ease with which an assessment tool or programme can be implemented by considering the resources required to deliver the assessments and the acceptability to the trainees.

Reliability
This is an expression of the consistency and reproducibility (precision) of measurements so that as far as possible sources of measurement errors have been taken into account. There are three important dimensions of error: the consistency of marking, the quality of the test and the test items themselves.

Sampling
In this context refers to gathering multiple ‘snapshots’ of the performance of the learner over the time-frame of training using multiple tools and incorporating the views of multiple individuals in order to build up a rich picture of the learner.

Supervision
Usually a formal one-to-one relationship, focused around professional conversations to help the supervisee develop reflective professional practice, learning and skills with the aim of improving patient care.
Triangulation
The principle that in judging the attainment of professional competence, a variety of tools are used to gather evidence, and that information gathered from one tool is compared and contrasted with evidence from other sources.

Utility
The usefulness of an assessment defined as a product of its reliability, validity, cost effectiveness, acceptability and educational impact.

Validity
Validity in assessment terms refers to the degree to which a measurement instrument truly measures what it is supposed to measure. It is concerned with whether the right things are being assessed, in the right way, and with a positive influence of learning.
Self-Assessment Activities

The following activities require you to review the programme of workplace-based assessment for your own specialty.

1. Competencies

Can you note down the competencies that are being assessed within this programme?

Do you know which tools are being used, and how to use them?

2. The impact of assessment on the learner

Next time you use one of these tools, what would you do differently to improve the educational impact of the assessment on the learner?

3. Quality assurance

Consider how you might quality assure this programme of assessment within your faculty, practice or department.

Identify three broad areas that you could improve, and for one of these areas identify some developmental outcomes.