



The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: An historical and theoretical perspective

Kamran Z. Khan, Sankaranarayanan Ramachandran, Kathryn Gaunt & Piyush Pushkar

To cite this article: Kamran Z. Khan, Sankaranarayanan Ramachandran, Kathryn Gaunt & Piyush Pushkar (2013) The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: An historical and theoretical perspective, Medical Teacher, 35:9, e1437-e1446, DOI: [10.3109/0142159X.2013.818634](https://doi.org/10.3109/0142159X.2013.818634)

To link to this article: <https://doi.org/10.3109/0142159X.2013.818634>



Published online: 22 Aug 2013.



Submit your article to this journal [↗](#)



Article views: 41382



View related articles [↗](#)



Citing articles: 111 View citing articles [↗](#)

WEB PAPER
AMEE GUIDE

The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: An historical and theoretical perspective

KAMRAN Z. KHAN¹, SANKARANARAYANAN RAMACHANDRAN², KATHRYN GAUNT² & PIYUSH PUSHKAR³

¹Manchester Medical School, UK, ²Lancashire Teaching Hospitals NHS Trust, UK, ³Aintree University Hospitals Trust, UK

Abstract

The Objective Structured Clinical Examination (OSCE) was first described by Harden in 1975 as an alternative to the existing methods of assessing clinical performance (Harden et al. 1975). The OSCE was designed to improve the validity and reliability of assessment of performance, which was previously assessed using the long case and short case examinations. Since then the use of the OSCE has become widespread within both undergraduate and postgraduate clinical education. We recognise that the introduction of the OSCE into an existing assessment programme is a challenging process requiring a considerable amount of theoretical and practical knowledge. The two parts of this Guide are designed to assist all those who intend implementing the OSCE into their assessment systems. Part I addresses the theoretical aspects of the OSCE, exploring its historical development, its place within the range of assessment tools and its core applications. Part II offers more practical information on the process of implementing an OSCE, including guidance on developing OSCE stations, choosing scoring rubrics, training examiners and standardised patients and managing quality assurance processes. Together we hope these two parts will act as a useful resource both for those choosing to implement the OSCE for the first time and also those wishing to quality assure their existing OSCE programme.

Introduction

Conducting an Objective Structured Clinical Examination (OSCE) for the first time is a complex and time-consuming task. It requires considerable understanding of the underlying educational principles of an OSCE and the development of academic and administrative structures to support and implement the examination. These supporting structures or entities could be in the form of teams and committees as described in detail in the second part of this Guide. In the institutions where the OSCE is already being used as an assessment tool, quality assurance and continuous improvement are important in order to maintain standards and psychometric rigour.

The current literature appears to be lacking a detailed and comprehensive manual to help institutions with the practicalities of implementing the OSCE for the first time, though a lot of studies covering various aspects of the OSCE have been published in peer reviewed journals. This Guide will present an evidence-based perspective on setting up an OSCE for those new to the approach, and will also provide some guidance and thought to those who would like to revisit their programmes for quality assurance purposes.

The Guide consists of two parts; Part I focuses on the historical background and educational principles of the OSCE. Knowledge and understanding of these principles is essential

Practice points

- Before the advent of OSCE, long case and short case examinations were used for the assessment of performance.
- The OSCE is an assessment tool based on the principles of objectivity and standardisation, which allows the assessment of candidates' performance against standardised scoring schemes by trained assessors.
- The OSCE assesses performance in a simulated environment, at the 'shows how' level of Miller's pyramid of assessment.
- The OSCE is most appropriately used, alongside other assessment methods, in a structured programme of assessment.
- A well-designed OSCE can drive learning, and therefore, can have a positive educational impact.

before moving any further in designing and administering an OSCE. We hope that the contents of Part I will act as a suitable and informative introduction for the readers, enabling them eventually to understand and implement the ideas and practical advice given in Part II, which will describe the organisation and administration of the OSCE.

Correspondence: Dr Kamran Khan, Department of Anaesthetic, Mafraq Hospital, Mafraq, Abu Dhabi, UAE; email: Kamran.Khan950@gmail.com

Assessment of performance and the development of the OSCE

Assessment of performance in health care education is dependent upon the choice of appropriate tools to measure the outcomes in question. According to Stokes (1979) such tools are designed for determining whether one can 'do things' rather than simply remember, talk and write about them. Any assessment tool used alone is not capable of assessing the combination of knowledge, skills and behaviours, which influence performance within the various contexts of health care. Epstein (2007) argues for the use of a combination of assessment tools to measure a range of learning domains; such a methodology can be called a 'test battery' approach (Hamdy et al. 2010). The OSCE plays an important part in this test battery in the assessment of performance in simulated environments (PMETB 2007; Khan & Ramachandran 2012).

Assessment in health care education has been constantly evolving; from the assessment of knowledge in the period up to the 1960s, to the assessment of performance from the 1970s onwards (Epstein 2007; Hays 2008). A range of assessment tools have subsequently been developed for the assessment of performance, including the OSCE with standardised patients, workplace based assessments and assessments using computer enhanced simulations (Norcini & McKinley 2007).

Prior to the development of the OSCE in the 1970s by Harden (1975), combinations of 'short case' and 'long case' examinations were used in an attempt to assess performance. Frequently, these methods also included an oral or viva voce examination (Sood 2001). These assessment tools are still being used widely in different parts of the world and warrant some consideration before we move on to discuss the OSCE.

Short case examinations

In short case examinations candidates are asked to perform a brief and focused clinical examination of around five or six patients with specific clinical findings. Fifteen to twenty minutes per patient are usually allocated. Following this, the findings and diagnosis are discussed with a pair of examiners. The examiners award marks independently on the candidates' performance; however, this scoring is often unstructured (Ponnamperuma et al. 2009), with the same examiners examining the candidate throughout the entire time.

Although the patients have real signs and symptoms, these patients often change between candidates, affecting the standardisation of the examination (Norcini 2002; Wass & van der Vleuten 2004). This format of examination has the advantage of having multiple, brief and observed real patient encounters similar to a clinicians' outpatient experience and are therefore considered as a very good assessment tool for clinical examination skills (Ponnamperuma et al. 2009). However, the reproducibility and validity of these examinations is affected by unstructured questioning by the examiners, a lack of standardisation of patients between candidates, and a lack of ability to assess history taking (Walsh 2006).

Viva voce examinations

In a typical Viva voce assessment, the candidates are provided with clinical material such as a case transcript, pathological specimen or investigation results. After a given time period to assimilate the information, two examiners question the candidates for 10–15 min. These examinations allow the assessment of recall, depth of knowledge, hypothetical problem solving, judgement, clinical reasoning and analytical skills (Jayawickramarajah 1985). Although typically unstructured questioning leads to poor inter-rater reliability (Harden et al. 1975); however, it is possible to standardise these items in order to minimise variance (Ponnamperuma et al. 2009). As there is no direct observation of the candidates' clinical performance the Viva voce can only assess this by proxy.

Long case examinations

During long case examinations, candidates are asked to take a history and perform a complete physical examination on a real patient, frequently chosen from the current in-patient or out-patient cohort. Candidates at different examination sittings may be allocated different patients with varying conditions and clinical signs. Candidates are asked to take a history and perform a complete physical examination in the first 30–45 min, often unobserved by the examiners (Ponnamperuma et al. 2009). Typically, unstructured questioning of the candidates follows this, which is usually focused on their clinical findings, diagnosis and management plan of the patients' problems. The candidates' interaction with the patient, including history taking, general communication and clinical skills is not always observed. Most frequently the discussion is based on the theoretical aspects of the case in question, exploring the depth of candidates' understanding i.e. their knowledge and the management plan for the patient (Sood 2001; Wass et al. 2001a; Wass & van der Vleuten 2004).

A number of drawbacks with this form of assessment have been identified; for example, the marks awarded for long cases are based upon unstructured questioning rather than checklists or standardised scoring sheets (Norman 2002). There is an inherent inter-case variability at separate sittings of the same examination (Norcini 2002; Wass & van der Vleuten 2004). Further, a single case does not allow the assessment of the candidates' performance on a range of areas, restricting the sampling of the 'test universe'. The content specificity of long cases makes the generalisation of the skills assessed using this technique nearly impossible (Wass et al. 2001b; Norcini 2002).

Despite the disadvantages described above, long case examinations have a high 'face validity', as the encounter between the candidate and the patient is realistic and tests the integrated interaction which is very close to actual daily practice (Wass & van der Vleuten 2004). Table 1 lists the advantages and disadvantages of long cases (Table 1).

To overcome the above mentioned drawbacks, many modifications of the long case examinations have been developed. The major focus of these modifications is on the observation of the candidates' performance during the long case. For instance, the Observed Long case (Newble 1991), the Leicester Assessment Package (Fraser et al. 1994), the Objective Structured Long Case Record (OSLER)

Table 1. Advantages and disadvantages of long cases.

Advantages of long cases	Disadvantages of long cases
The use of real patients achieves the highest degree of realism	Inter-case variance or case specificity (Wass et al. 2001b)
Ability to assess on rare or complex problems with good clinical signs and symptoms in patients (Smith et al. 1984)	Real patient variance or inconsistency in case presentation between candidates (Wass et al. 2001b)
Ability to choose from a variety of cases (Smith et al. 1984)	Single case limiting the ability to assess a broad spectrum of skills
Complete evaluation of the patient's problem—no compartmentalisation	Undue leniency or stringency of the examiners based on the candidate or patient characteristics (McManus et al. 2006)
	Low inter-rater and intra-rater reliability (Wilson 1969; Norcini 2002)
	High case-specificity (Norman 2002)
	Non-standardised scoring (McManus et al. 2006)
	No blueprinting resulting in random selection of cases for assessment

Table 2. Assessment domains in the OSLER (Gleeson 1997).

Assessment domains	Specific criteria in each domain
History	Pace and clarity of presentation Communication skills Systematic approach Establishment of case facts
Physical examination	Systematic approach Examination technique Establishment of correct physical findings
Construction of appropriate investigations in a logical sequence	
Appropriate management	
Final clinical acumen (ability to identify and solve problems)	

(Gleeson 1997), the Standardised Structured Long Case Examination of Clinical Competence (Troncon et al. 2000) and the Direct Observation Clinical Encounter Examination (DOCEE) (Hamdy et al. 2003) are described in the literature.

Amongst all of the above, the OSLER has gained a degree of popularity. The reliability and construct validity of the OSLER is better than the long case; this has been achieved by the introduction of different domains to assess the performance of the candidates in each case (Gleeson 1997) (Table 2).

In order to achieve the high levels of reliability demonstrated in the literature (Cronbach's alpha of 0.8), observed long case examinations require long testing times of up to three and a half hours which may be impractical at some institutions (Wass et al. 2001a). The concepts of reliability and validity are described in some detail later.

In short, the main criticisms of the traditional methods of assessing clinical performance are a lack of structure and standardisation. The development of the OSCE by Harden (1975) aimed to address the above issues and improve the quality of the assessment of performance. At the time of its inception, it was described as a tool for the assessment of competence but more recently competence is seen as a point on the spectrum of performance. By this inference the OSCE is a tool used for the assessment of performance in simulated environments (Khan & Ramachandran 2012); this concept is explored in some detail later in this Guide.

The original OSCE

Prior to the development of the OSCE, the candidate's assessment could be affected by the patient's performance, examiner bias, a non-standardised marking scheme and the candidate's actual performance. The OSCE was first designed to introduce standardisation and reduce the number of variables that could impact the assessment of performance. Hence, in a well-designed OSCE, the grades of the candidates should predominantly be affected by the performance of the candidates alone, with minimal effect from other sources of variance. In addition, the OSCE was designed as a novel evaluation tool, allowing the assessment of candidates' clinical skills, attitudes, problem-solving abilities, and their application of knowledge in one examination (Harden et al. 1975).

The first OSCE was conducted by Harden in 1972 in Dundee, and described in the literature in 1975 (Harden et al. 1975). Since then a healthy body of published literature has developed to support and substantiate an evidence-based use of OSCE techniques, underpinned by sound educational principles (Carraccio & Englander 2000; Wass et al. 2001b; Hodges et al. 2002; Turner & Dankoski 2008; Boursicot 2010; Gupta et al. 2010; Zayyan 2011).

In the original version described by Harden, the students moved around 18 testing and 2 rest stations in a hospital ward. Each station was 4.5 min long with a 30 s break between the stations. The total examination time was 100 min. Each station tested a single competency; for instance performance of a procedure, history taking or clinical examination of a patient. The original 18 stations and the sequence as described by Harden (1979) are shown (Figure 1). The standardised marking scheme used at station 3 in this examination is also reproduced in the subsequent figure (Figure 2).

Defining the OSCE

Since the original OSCE, many definitions of this assessment technique have been proposed; Harden (1988) defined it as; "An approach to the assessment of clinical competence in which the components are assessed in a planned or structured way with attention being paid to the objectivity of the examination".

According to Newble (2004) "The OSCE is not a test method in the same way as an essay or multiple-choice question. It is basically an organization framework consisting

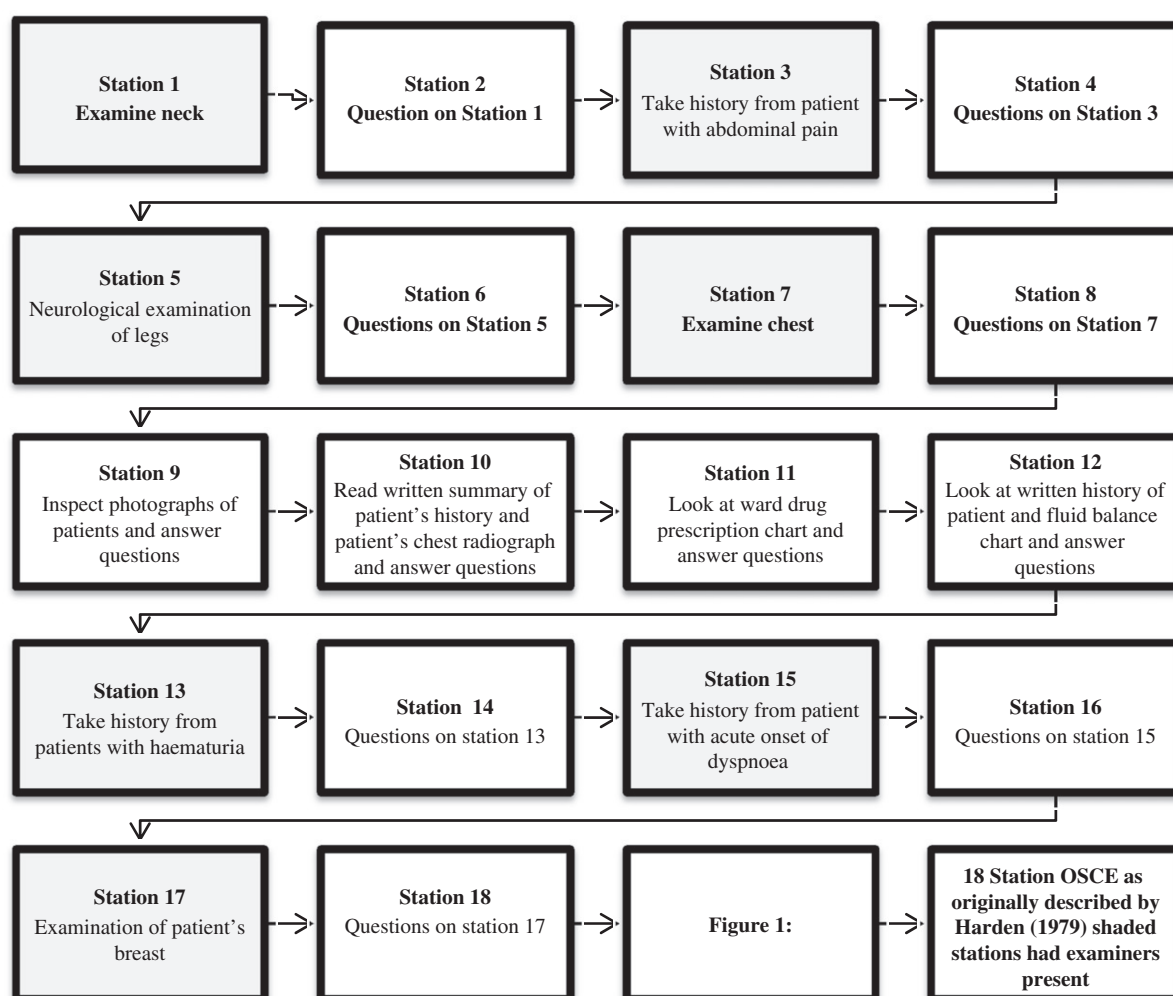


Figure 1. Eighteen station OSCE as originally described by Harden (1979). Shaded stations had examiners present.

of multiple stations around which students rotate and at which students perform and are assessed on specific tasks". Hodder (1989) and van der Vleuten and Swanson (1990) support this view that many different types of test methods can be incorporated within the OSCE format.

Based on various descriptions of OSCE in the literature we propose a consolidated definition of the OSCE; "An assessment tool based on the principles of objectivity and standardisation, in which the candidates move through a series of time-limited stations in a circuit for the purposes of assessment of professional performance in a simulated environment. At each station candidates are assessed and marked against standardised scoring rubrics by trained assessors".

Some variants of the OSCE described in Table 3 use the original format of moving around assessment stations to assess a range of different outcomes (Table 3).

What Does the OSCE Assess?

It is important to understand the relationship of competence and performance before exploring the details of what an OSCE can be used to assess. The OSCE has previously been described in the literature as a tool used for the assessment of competence. This is slightly misleading as according to the

Dreyfus & Dreyfus (1980) competence is a point on the spectrum of performance. Any tool used for the assessment of performance should be able to grade the candidates as novice, advanced beginner, competent, proficient or experts in the tasks allocated to them. A key difference between the OSCE and Work Place Based Assessment (WPBA) is the setting in which performance is being assessed, not that the former assesses competence and the latter performance, as is commonly perceived (Boursicot et al. 2011). Nevertheless, difference is very significant since the performance of an individual on identical tasks can vary considerably depending on the context of the assessment (ten Cate et al. 2010). Therefore, for all practical purposes the OSCE should be seen as a tool for the assessment of performance within simulated environments. Competency is the composite of cognitive, psychomotor and affective skills as appropriate, while competence is an attribute of a person. Assessment of performance on competencies, for example; the ability to communicate or examine the respiratory system allows the assessor to determine if the performer is competent, proficient or expert, etc. A detailed discussion on this topic is beyond the scope of this Guide and we have addressed it in our paper on competency, competence and performance (Khan & Ramachandran 2012).

Student's name and initials:

Instructions to student:

This patient has been admitted for investigation of abdominal pain.

Obtain a history

A. General patient information and key points in the history: Please tick the appropriate boxes. Where there is no tick it will be assumed that the student did not attempt the procedure.	Carried out satisfactorily	Attempted but not satisfactorily
Occupation		
Pain: Type		
Site		
Radiation		
Relieving factors		
Exacerbating factors		
Periodicity		
Duration		
Appetite		
Nausea, vomiting		
Weight loss		
Bowel habit		
Blood in stools		
Family history		
Drug history		
Smoking		
Alcohol		
Previous medical history, e.g. per haematemesis		

B. History taking technique

Please mark out of 10 the student's history taking technique, taking account of:

- dates established
- correct phrasing of questions
- attention paid to answers
- answers followed up appropriately
- systematic approach
- effective use of time

- 8–10 Distinction
7– Very good pass
6– Pass
5– Bare pass
4– Fail
3 or less Bad fail

 B

C. Student's attitude to patient

Please mark out of 10 the student's relationship with the patient, taking account of:

- no unnecessary discomfort to patient
- consideration of patient's feelings
- attempt to establish rapport with patient

- 8–10 Distinction
7– Very good pass
6– Pass
5– Bare pass
4– Fail
3 or less Bad fail

 C

 A

Figure 2. Scoring rubric used at station 3 as described by Harden (1979).

The OSCE and its variants described earlier are only a few from a number of tools available to assess performance in health care. The diagram (Figure 3) describes the use of Workplace Based Assessments and Incognito Patients (Rethans et al. 2007) as additional methods for the

competency based assessment of performance (CBAP), in different settings.

Figure 3 highlights that assessment at an OSCE station gives a 'snapshot' of candidates' demonstrated performance in a particular area in a simulated environment. With reference to

Table 3. Variations of the OSCE.

OSCE variant	Brief description
Objective Structured Practical Examination (OSPE) (Harden & Cairncross 1980)	Assessment of practical skills, knowledge and/or interpretation of data in non clinical settings
Objective Structured Assessment of Technical Skills (OSATS) (Bodle et al. 2008; van Hove et al. 2010)	Designed for objective skills assessment, consisting of a global rating scale and a procedure specific checklist. It is primarily used for feedback or measuring progress of training in surgical specialities.
Objective Structured Video Examinations (OSVE) (Humphris & Kaney 2000; Vantis et al. 2004)	Videotaped recordings of patient-doctor encounters are shown to students simultaneously and questions related to the video clip are asked. Written answers are marked in a standardised manner.
Team Objective Structured Clinical Examination (TOSCE) (Singleton et al. 1999)	Formative assessment covering common consultations in general practice. A team of students visits each station in a group, performing one task each in a sequence. The candidates are marked for their performance and feedback is provided. The team approach improves efficiency and encourages learning from peers.

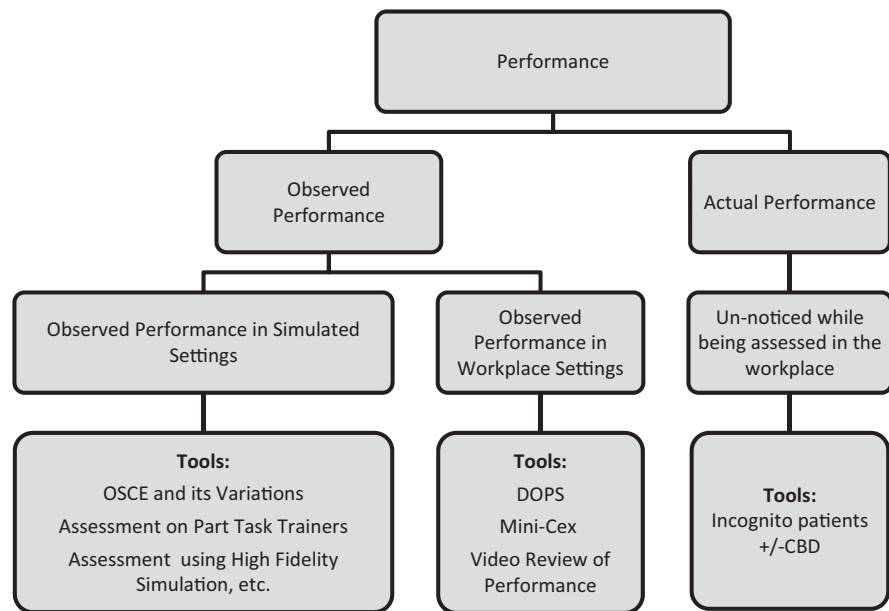


Figure 3. Model for the competency based assessment of performance and examples of available assessment tools. Reproduced from Khan & Ramachandran (2012).

Miller's pyramid of assessment (Miller 1990), the OSCE assesses that a candidate is able to 'show how' one would perform in a simulated environment (Figure 4). This contrasts to the other methods, such as 'incognito patients', which allows assessment of the candidates' actual performance in their workplace, at the 'does' level of the pyramid.

The simulated environment itself can influence the performance of candidates; therefore, the performance of the candidates in the OSCE might not be the same as their performance in the workplace on identical tasks (Khan & Ramachandran 2012), hence it is important to bear this in mind when interpreting the results. A number of other factors also influence performance in the health care setting as outlined in Figure 5.

In real life, non-clinical skills such as team working, resource management, situational awareness and leadership, etc. also play a very important part in determining the overall performance. It is not very easy to assess these using the OSCE

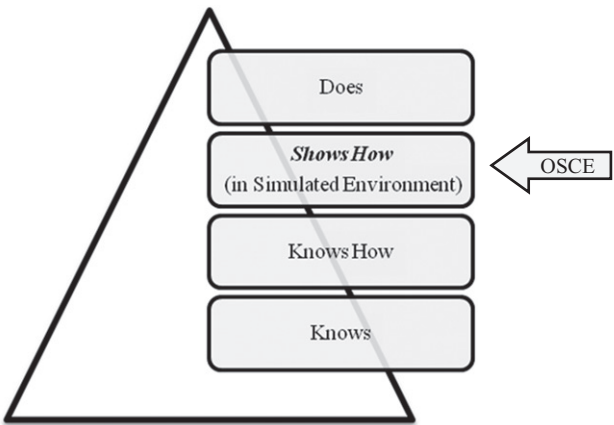


Figure 4. The OSCE in relation to the Miller's pyramid.

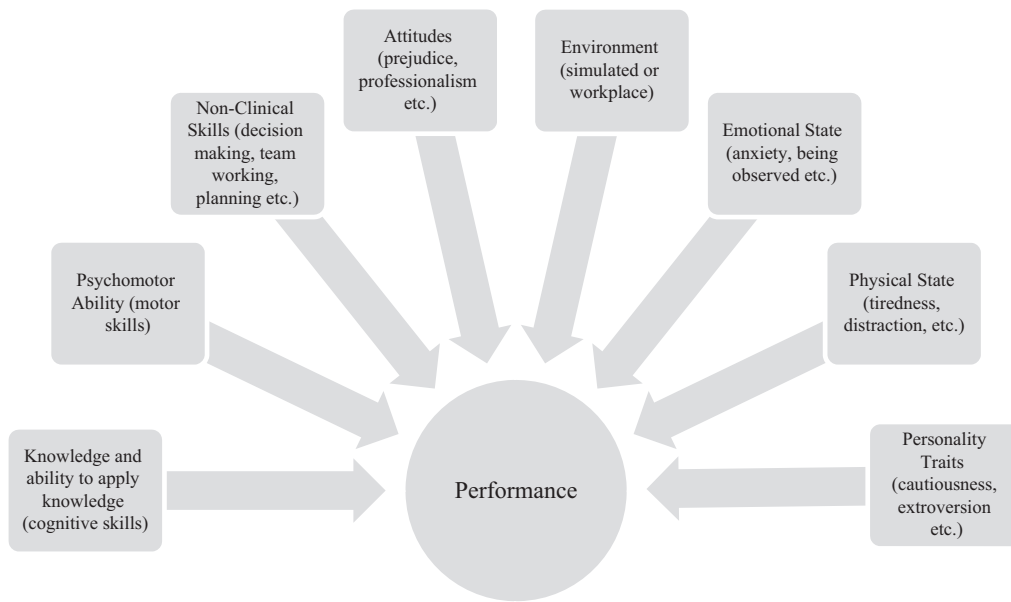


Figure 5. Factors influencing performance. Modified from Khan & Ramachandran (2012).

Table 4. Affective, cognitive and psychomotor domains of learning from basic to advanced levels.

Range of difficulty	Affective	Cognitive	Psychomotor
Basic	Receiving Responding Valuing Organisation Characterisation	Knowledge Comprehension Application Analysis Synthesis Evaluation	Perceiving Patterning Accommodation Refining Improving Composing
Advanced			

format. Assessment using the OSCE mainly focuses on cognitive, psychomotor and affective skills, described as learning domains. Krathwohl, Bloom and Jewett have described taxonomies of learning in affective, cognitive and psychomotor domains, respectively; each with a range of levels of difficulty. These levels are shown for each of the three learning domains in Table 4 (Krathwohl 1956; Jewett et al. 1971; Bloom 1974).

It is possible for the OSCE to assess skills within all three learning domains; however, it should be borne in mind that there might be more appropriate methods of assessing certain skills. For instance, the theoretical application of knowledge is better tested by using multiple choice questions (MCQs), and evaluation of knowledge can be assessed better by critical appraisal of a scholarly article. As previously described, an ideal summative assessment programme needs to be based on a combination of assessment tools in a 'test battery approach', thus allowing the assessment of cognitive, affective and psychomotor skills, mapped to a range of learning outcomes. It is also important to consider that it is not possible to test all levels of difficulty within each domain with the use of an OSCE, for instance it would be very difficult to plan a station testing the ability of the candidates to 'evaluate knowledge' in the cognitive domain.

We believe that, the OSCE should be designed to assess certain competencies or skills which are not possible to be assessed using pen and paper or computer based testing methods. Typical examples of such skills could be the ability of a candidate to take a history or perform a procedure. However, there have been criticisms that the OSCE risks the compartmentalisation of these skills (performance of skills in isolation) for the purposes of assessment (Nestel et al. 2011). This results in a loss of ability to assess candidates' performance holistically. For this reason it is important that performance related skills are not tested in isolation, rather these are blended with skills such as the application of knowledge or formulating a management plan so that performance is assessed more holistically.

Consider stations designed to assess candidates' ability to interpret an X-ray or to recognise a raised urea and list possible causes; an OSCE station is not required for such assessments. In fact, paper-based methods could assess these cognitive skills reliably, and prove to be less resource intensive. Also consider a station designed to assess candidates' ability to perform the psychomotor component of a 'pap smear test' on a model; such a station risks compartmentalisation as in reality candidates would need to communicate with the patient and apply their knowledge on this procedure whilst they perform it. Poorly structured OSCE stations may lead to candidates learning skills to pass examinations rather than to improve their actual clinical performance. At the same time it is also important to prevent assessment overload by trying to assess too many skill subsets of performance at a single station. The example of such a scenario could be asking the candidates to take a history and examine a patient complaining of haemoptysis, interpret their X-ray findings, and explain the likely diagnosis to the patient. Although this station may reflect 'real life' by placing skills within context it is unlikely that assessing all of these skills adequately within the one station is achievable within time constraints.

Educational principles of the OSCE

The two major underlying principles of the OSCE are 'objectivity' and 'structure'. Objectivity predominantly depends on standardised scoring rubrics and the same, trained, examiner asking the same questions to every candidate. A well-structured OSCE station on the other hand has a standardised station design assessing a specific clinical task which is blueprinted against the curriculum. A well-designed OSCE has a high level of *validity* (Downing 2003), which in simple terms means that the OSCE assesses what it is designed to assess. At the same time well-designed OSCE also has been shown to demonstrate a high degree of *reliability* (Boursicot 2010), i.e. the examination results are reproducible with very little error. The concepts of validity, reliability, feasibility, and educational impact are explored in more detail below.

Validity

Previously different types of validity such as face validity, content validity, and construct validity, etc. were seen as separate entities. More contemporaneously, construct validity is seen as an overarching term covering all types of validity (Downing 2010). By this new definition, reliability itself is also seen as an element of construct validity, which was previously seen as a separate entity. It is beyond the scope of this Guide to discuss this in further detail but generally construct validity depends on five sources of evidence (American Educational Research Association 1999):

- (1) The test content represents what the curriculum needs to assess, the tasks are realistic and the right domains are being assessed.
- (2) The responses to the test item are accurately recorded, handled, stored, and analysed.
- (3) The test has high reliability.
- (4) The test results correlate with other test results, assessing similar domains and show poor correlation with those assessing different domains (convergence and divergence).
- (5) The consequences of assessment are sound, i.e. effect on the learning is positive, the penalties for failure are justifiable and the consequences of passing are socially and professionally acceptable.

The above principles are applicable to the OSCE and these are discussed in the second part of this Guide.

Reliability

Many factors influence the reliability of any assessment method and an understanding of these is essential when developing a new assessment programme. The main influences on the reliability of an OSCE are outlined briefly below; each of these areas is discussed in considerable detail in the relevant sections in the second part of this Guide.

(1) *The number of stations*

Testing candidates across a large sample of clinical cases maximises reliability (Roberts et al. 2006) and an appropriate

test length has been shown to have the greatest influence in ensuring that candidates' overall performance is reliably assessed (Swanson et al. 1995).

(1) *Standardised scoring rubrics*

These ensure that examiners are marking candidates against the same criteria thereby improving consistency of scoring between candidates and examiners (Smee 2003).

(1) *Using trained examiners*

There is some evidence suggesting that examiner training reduces examiner variation in scoring (Newble et al. 1980; van der Vleuten et al. 1989) and improves consistency of examiners' behaviour. Further using different examiners for different stations can reduce individual assessor bias (Gormley 2011).

(1) *Standardised patient performance*

Poorly standardised patients who vary their performance between candidates can reduce the reliability of the examination (Smee 2003).

Feasibility

Feasibility can be defined as the degree of practicality of the assessment intervention. Compared to other assessment methods, the OSCE is seen as more resource intensive and time-consuming to set-up and run. For this reason it is important to ensure that the best use is made of the OSCE by only developing stations that require an OSCE format, i.e. assessment of performance rather than knowledge.

Educational impact

The OSCE can have a *positive educational impact* as it can drive learning (Boursicot 2010). This positive educational impact is dependent on realistic recreation of assessment scenarios at the OSCE stations. If candidates find it difficult to differentiate between the assessment tasks and real life practice then the OSCEs can drive lifelong learning. If the tasks given at OSCE stations are compartmentalised and driven by checklist scoring then the candidates learn to pass exams, decreasing the educational impact of the OSCE (Miller 1990; Shumway & Harden 2003).

Common uses of the OSCE

In the last three decades the OSCE has seen a steady exponential growth and usage in both undergraduate and postgraduate examinations around the globe. The OSCE is also used for licensure examinations and as a feedback tool in formative settings. Common uses of the OSCE are listed below.

- (1) As a performance based assessment tool for testing the minimum accepted standards of students or trainees as barrier (exit) examinations during the undergraduate years in most of the medical schools in the United States, United Kingdom and Canada.
- (2) As a postgraduate high stakes assessment tool in Royal College examinations, e.g. the Royal College of Physicians (United Kingdom) uses the Practical

Assessment of Clinical Examination Skills (PACES) examination as a component of its membership examinations (Wallace 2002).

- (3) As a formative assessment tool in undergraduate medical education (Townsend et al. 2001).
- (4) As a tool for the assessment of graduates seeking high-stakes licensure and certification to practise medicine (Hodges 2003). Professional Linguistics and Assessment Board Part II [PLAB] in United Kingdom (Tombleson et al. 2000), Medical Council of Canada Qualifying Examination II (Reznick et al. 1992), and the Clinical Skills Assessment part of the United States Medical Licensure Examination are based on an OSCE model (Whelan 1999).
- (5) As an educational tool to provide immediate feedback (Hodder et al. 1989; Brazeau et al. 2002).

Conclusions

Part I of this Guide has introduced the concept of the OSCE as a tool for performance-based assessment in simulated environments. This part has also described why the OSCE was developed and how it can meet the need for more structure and objectivity within the assessment of performance. We have stressed throughout this Guide that the OSCE should be used only when other methods of assessment cannot assess the competency in question or in conjunction with other assessments of performance. It neither replaces the tools used for assessment of knowledge nor is it a panacea for all the assessment needs in an educational programme. The challenges associated with the academic and practical aspects of implementing an OSCE, particularly when this examination format is novel to the institution are addressed in the Part II of this Guide. Readers equipped with the information in this part of the Guide will find it easy to understand the complex processes needed for the setup of OSCE programmes, described in the second part.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

Notes on contributors

Dr KAMRAN KHAN, MBBS, FRCA, FAcadMed, MScMedEd, at the time of production of this manuscript was Senior Teaching Fellow Manchester Medical School and Consultant Anaesthetist, LTHTR, Preston, UK. During his career Dr Khan has developed a passion for medical education and acquired higher qualification in this field alongside his parent speciality in Anaesthetics. He has held clinical academic posts, first at the University of Oxford and then at the University of Manchester. He was Associate lead for 'Assessments' at the Manchester Medical School. He is currently working in the UAE as a Consultant Anaesthetist. He has extensively presented and published at the national and international levels in his fields of academic interest.

Dr SANKARANARAYANAN RAMACHANDRAN, MBBS, Pg Cert (MedEd), FHEA, is a Medical Education & Simulation Fellow, Lancashire Teaching Hospitals NHS Trust, Preston, UK. He qualified in India and underwent postgraduate training in General Medicine. He had been teaching and training undergraduate and postgraduate medical students throughout his career. He has completed a Postgraduate Certificate and pursuing a Master's degree in Medical Education at the University of Manchester.

He has special interest in assessment, problem-based learning and research in medical education.

DR KATHRYN GAUNT, MBChB, MRCP, PGD(MedEd), at the time of production of this manuscript was a Medical Education and Simulation Fellow, LTHTR, Preston, UK. Dr Kathryn Gaunt graduated in 2002 from the University of Manchester Medical School, UK. She was a Specialty Registrar in Palliative Medicine and has a keen interest in medical education. She is currently pursuing a higher qualification in this field. She has recently taken time out of her training programme to work as a Medical Education and Simulation Fellow at the Lancashire Teaching Hospitals NHS Trust. Here she lectured, examined for OSCEs and was involved in research and development within the field. She has now returned to her training in Palliative Medicine.

Dr PIYUSH PUSHKAR, MBChB, works in Critical Care at the Aintree University Hospitals Trust, Liverpool, UK. Piyush qualified from the University of Edinburgh and has subsequently worked and trained in anaesthesia and critical care in Lancashire and Merseyside, as well as aeromedical retrieval medicine in Queensland, Australia. He also works as a volunteer doctor for Freedom from Torture in Manchester.

References

- American Educational Research Association. 1999. Standards for educational and psychological testing. Washington, DC: American Educational Research Association.
- Bloom BS. 1974. Taxonomy of educational objectives: The classification of educational goals handbook 1: Cognitive domain, [S.I.], New York: Longman.
- Bodde JF, Kaufmann SJ, Bisson D, Nathanson B, Binney DM. 2008. Value and face validity of objective structured assessment of technical skills (OSATS) for work based assessment of surgical skills in obstetrics and gynaecology. *Med Teach* 30:212–216.
- Boursicot K. 2010. Structured assessments of clinical competence. *Br J Hosp Med* 71:342–344.
- Boursicot K, Etheridge L, Setna Z, Sturrock A, Ker J, Smee S, Sambandam E. 2011. Performance in assessment: Consensus statement and recommendations from the Ottawa conference. *Med Teach* 33:370–383.
- Brazeau C, Boyd L, Crosson J. 2002. Changing an Existing OSCE to a Teaching Tool: The Making of a Teaching OSCE. *Acad Med* 77:932.
- Carraccio C, Englander R. 2000. The Objective Structured Clinical Examination. A step in the direction of competency-based evaluation. *Arch Pediatr Adol Med* 154:736–741.
- Downing MS. 2003. Validity: On the meaningful interpretation of assessment data. *Med Educ* 37:830–837.
- Downing SM. 2010. Validity: Establishing meaning for assessment data through scientific evidence. St George's Advanced Assessment Course, London.
- Dreyfus S, Dreyfus H. 1980. A Five stage model of the mental activities involved in directed skill acquisition. Research Paper. Department of Philosophy, University of California, Berkeley.
- Epstein RM. 2007. Assessment in Medical Education. *N Eng J Med* 356:387–396.
- Fraser RC, Mckinley RK, Mulholland H. 1994. Consultation competence in general practice: Establishing the face validity of prioritised criteria in the Leicester assessment package. *Br J Gen Pract* 44:109–113.
- 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.
- Gormley G. 2011. Summative OSCEs in undergraduate medical education. *Ulster Med J* 80:127–132.
- Gupta P, Dewan P, Singh T. 2010. Objective Structured Clinical Examination (OSCE) Revisited. *Indian Pediatr* 47:911–920.
- Hamdy H, Prasad K, Williams R, Salih FA. 2003. Reliability and validity of the direct observation clinical encounter examination (DOCEE). *Med Educ* 37:205–212.
- Hamdy H, Telmesani AW, Wardy NA, Abdel-Khalek N, Carruthers G, Hassan F, Kassab S, ABU-Hijleh M, AL-Roomi K, O'malley K, et al. 2010. Undergraduate medical education in the Gulf Cooperation Council: A multi-countries study (Part 2). *Med Teach* 32:290–295.

- Harden RM. 1979. How to Assess Clinical Competence – An overview. *Med Teach* 1:289–296.
- Harden RM. 1988. What is an OSCE? *Med Teach* 10:19–23.
- Harden RM, Cairncross RG. 1980. Assessment Of Practical Skills: The Objective Structured Practical Examination (OSPE). *Studies High Educ* 5:187–196.
- Harden RM, Stevenson M, Downie WW, Wilson GM. 1975. Assessment of Clinical Competence using Objective Structured Examination. *BMJ* 1:447–451.
- Hays R. 2008. Assessment in medical education: Roles for clinical teachers. *Clin Teach* 5:23–27.
- Hodder RV, Rivington RN, Calcutt LE, Hart IR. 1989. The effectiveness of immediate feedback during the Objective Structured Clinical Examination. *Med Educ* 23:184–188.
- Hodges B. 2003. OSCE! Variations on a theme by Harden. *Med Educ* 37:1134–1140.
- Hodges B, Hanson M, Mchaughton N, Regehr G. 2002. Creating, Monitoring, and Improving a Psychiatry OSCE. *Acad Psychiat* 26(3):134–161.
- Humphris GM, Kaney S. 2000. The Objective Structured Video Exam for assessment of communication skills. *Med Educ* 34:939–945.
- Jayawickramarajah PT. 1985. Oral examinations in medical education. *Med Educ* 19:290–293.
- Jewett AE, Jones SL, Luneke SM, Robinson SM. 1971. Educational change through a taxonomy for writing physical education objectives. *Quest* 15:32–38.
- Khan K, Ramachandran S. 2012. Conceptual Framework for Performance Assessment: Competency, Competence and Performance in the Context of Assessments in Healthcare – Deciphering the Terminology. *Med Teach* 34:920–928.
- Krathwohl DR. 1956. Taxonomy of educational objectives: The classification of educational goals; handbook 2: Affective domain, [S.l.], New York: Longmans.
- Mcmanus IC, Thompson M, Mollon J. 2006. Assessment of examiner leniency and stringency ('hawk-dove effect') in the MRCP(UK) clinical examination (PACES) using multi-facet Rasch modelling. *BMC Medical Education* 6:42.
- Miller GE. 1990. The assessment of clinical skills/competence/performance. *Acad Med* 65:S63–S67.
- Nestel D, Kneebone R, Nolan C, Akhtar K, Darzi A. 2011. Formative Assessment of Procedural Skills: Students' Responses to the Objective Structured Clinical Examination and the Integrated Performance Procedural Instrument. *Assess Eval Higher Educ* 36:171–183.
- Newble D. 2004. Techniques for measuring clinical competence: Objective structured clinical examinations. *Med Educ* 38:199–203.
- Newble DI. 1991. The observed long-case in clinical assessment. *Med Educ* 25:369–373.
- Newble DI, Hoare J, Sheldrake PF. 1980. The selection and training of examiners for clinical examinations. *Med Educ* 14:345–349.
- Norcini JJ. 2002. Death of a long case. *BMJ* 324:408–409.
- Norcini JJ, Mckinley DW. 2007. Assessment methods in medical education. *Teaching Teach Educ* 23:239–250.
- Norman G. 2002. The long case versus objective structured clinical examinations. *BMJ* 324:748–749.
- PMETB. 2007. Developing and maintaining an assessment system - a PMETB guide to good practice [Online]. London: PMETB. [Accessed 10 June 2012] Available from http://www.gmc-uk.org/assessment_good_practice_v0207.pdf_31385949.pdf
- Ponnampuruma GG, Karunathilake IM, Mcalder S, Davis MH. 2009. The long case and its modifications: A literature review. *Med Educ* 43:936–941.
- Rethans J-J, Gorter S, Bokken L, Morrison L. 2007. Unannounced standardised patients in real practice: A systematic literature review. *Med Educ* 41:537–549.
- Reznick R, Smee S, Rothman A, Chalmers A, Swanson D, Dufresne L, Lacombe G, Baumber J, Poldre P, Levasseur L. 1992. An objective structured clinical examination for the licentiate: Report of the pilot project of the Medical Council of Canada. *Acad Med* 67:487–494.
- Roberts C, Newble D, Jolly B, Reed M, Hampton K. 2006. Assuring the quality of high-stakes undergraduate assessments of clinical competence. *Med Teach* 28:535–543.
- Shumway JM, Harden RM. 2003. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med Teach* 25:569–584.
- Singleton A, Smith F, Harris T, Ross-Harper R, Hilton S. 1999. An evaluation of the Team Objective Structured Clinical Examination (TOSCE). *Med Educ* 33:34–41.
- Smee S. 2003. ABC of learning and teaching in medicine 'Skill based assessment'. *BMJ* 326:703–706.
- Smith IJ, Price DA, Houston IB. 1984. Objective structured clinical examination compared with other forms of student assessment. *Arch Dis Child* 59:1173–1176.
- Sood R. 2001. Long Case Examination - Can it be Improved. *Indian Acad Clin Med* 2:251–255.
- Stokes JF. 1979. Take a Clinical Examination. *BMJ* 1:98–99.
- Swanson DB, Norman GR, Linn RL. 1995. Performance-Based Assessment: Lessons From the Health Professions. *Educ Res* 24:5–11.
- Ten Cate TJO, Snell L, Carraccio C. 2010. Medical competence: The interplay between individual ability and the health care environment. *Med Teach* 32:669–675.
- Tombleson P, Fox RA, Dacre JA. 2000. Defining the content for the OSCE component of the PLABoard examination development of a blueprint. *Med Educ* 34:566–572.
- Townsend AH, McIlvenny S, Miller CJ, Dunn EV. 2001. The use of an objective structured clinical examination (OSCE) for formative and summative assessment in a general practice clinical attachment and its relationship to final medical school examination performance. *Med Educ* 35:841–846.
- Troncon LA, Dantas RO, Figueiredo JFC, Ferrioli E, Moriguti JLC, Martinelli ALC, Voltarelli JLC. 2000. A standardized structured long case examination of clinical competence of senior medical students. *Med Teach* 22:380–385.
- Turner JL, Dankoski ME. 2008. Objective Structured Clinical Exams: A Critical Review. *Fam Med* 40:574–578.
- Van Der Vleuten CPM, Swanson DB. 1990. Assessment of clinical skills with standardized patients: State of the art. *Teach Learn Med* 2:58–76.
- Van Der Vleuten CPM, Van Luyk SJ, Van Ballegooijen AMJ, Swansons DB. 1989. Training and experience of examiners. *Med Educ* 23:290–296.
- Van Hove PD, Tuijthof GJM, Verdaasdonk EGG, Stassen LPS, Dankelman J. 2010. Objective assessment of technical surgical skills. *Br J Surg* 97:972–987.
- Vlantis AC, Lee WC, Van Hasselt CA. 2004. The objective structured video examination of medical students. *Med Educ* 38:1181–1182.
- Wallace J. 2002. Simulated patients and objective structured clinical examinations: Review of their use in medical education. *Adv Psychiat Treat* 8:342–348.
- Walsh, K. 2006. OSCE. *BMJ Learning*. Available from group.bmj.com/products/learning/medical-presentations-2/OSCEs.ppt [Accessed 13 August 2013].
- Wass V, Jones R, Van Der Vleuten C. 2001a. Standardized or real patients to test clinical competence? The long case revisited. *Med Educ* 35:321–325.
- Wass V, Van Der Vleuten C. 2004. The long case. *Med Educ* 38:1176–1180.
- Wass V, Van Der Vleuten C, Shatzer J, Jones R. 2001b. Assessment of clinical competence. *Lancet* 357:945–949.
- Whelan GP. 1999. Educational commission for Foreign Medical Graduates - clinical skills assessment prototype. *Med Teach* 21:156–160.
- Wilson GM. 1969. Examination of clinical examiners. *Lancet* 4:37–40.
- Zayyan M. 2011. Objective Structured Clinical Examination: The Assessment of Choice. *Oman Med J* 26(4):219–222.