An introduction to assessing clinical skills

Sophie Hill[1] and Rich Bregazzi[2]

1. Lecturer in Medical Education, School of Medicine, Newcastle University, UK
2. Lecturer in Medical Education, Newcastle University, UK; Visiting Research Fellow in Healthcare Education, St John's College, Durham University, UK

ABSTRACT

The successful acquisition of clinical skills is essential to development and competence as a clinician. Clinical skills can be assessed in undergraduate education and in the workplace after graduation. Clarity about what is being assessed, and why, should support the development of any assessment process.

Keywords: Clinical skills assessment, workplace-based assessment, objective structured clinical examination.

INTRODUCTION

The assessment of clinical skills occurs in various formats. At medical school we often assess specific skills, for example being able to insert a cannula, using an Objective Structured Clinical Examination (OSCE). In the clinical workplace we assess more complex competencies that require a combination of skills and behaviours. These Workplace Based Assessments (WBA) involve supervisors and other members of the clinical team and use a variety of assessment formats. In this article we introduce some of the principles and approaches to the assessment of clinical skills.

PURPOSE AND PRINCIPLES OF ASSESSMENT

It is helpful for both teachers and learners to understand the purpose and the principles of assessment.

Our initial thoughts should consider what we want to assess, and why. Does an assessment identify current strengths and areas for development, to encourage learning? Or is it to make a judgement on an individual's knowledge or clinical competence at a specific point in time? The first is formative assessment, assessment for learning. A previous article[1] provided suggestions for how you might approach this. The second is summative and is assessment of learning. Summative assessment may include a formative element, by providing feedback to the learner, but this is not always possible.

Whatever the purpose of the assessment, it should be designed to enable the learner to demonstrate that they have achieved the intended learning outcomes. For example, can a learner describe the anatomy of the shoulder, or can they correctly insert a cannula? The approach taken to learning and teaching, the intended learning outcomes, and the assessment process, must align with each other. This is known as constructive alignment.[2]

There are a number of taxonomies available that help us understand the level and complexity of what is required of a performance, and what we intend to assess. Bloom's[3] cognitive taxonomy is often used for writing learning outcomes and examination questions. For assessment of clinical competence Bloom's[3] affective and psychomotor taxonomies may be more appropriate. In addition, there is Miller's[4] pyramid (Figure 1) where learners progress from novice to expert through four levels.
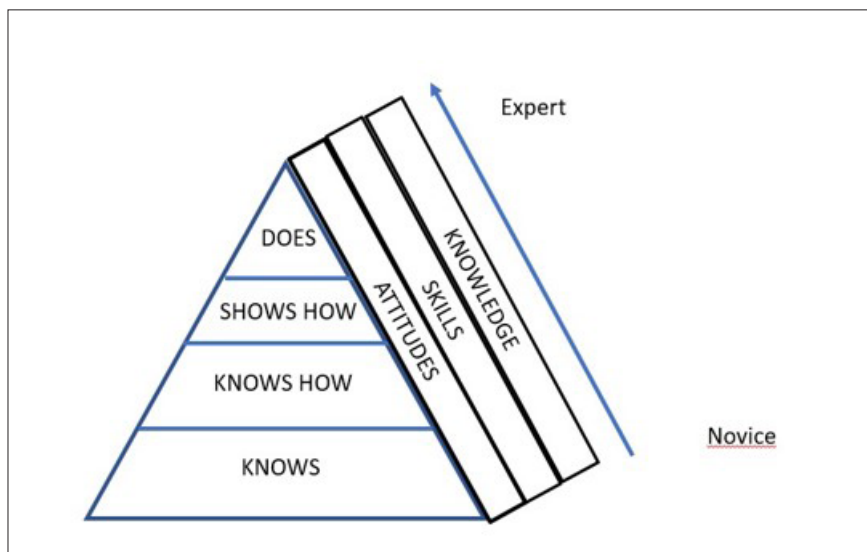
Figure 1. Miller's pyramid[4]

The lower two levels, 'knows' and 'knows how' concern assessment of knowledge; the top two, 'shows how' and 'does', are appropriate for clinical skills assessments.

'Shows how' demonstrates that a learner can carry out a task. This can be assessed using an Objective Structured Clinical Examination (OSCE), a simulation, or a case study. 'Does' involves real life performance. Here, assessment requires observation in practice, often using work-based assessment documents to record the outcome.

Wass and Archer[5] expand Miller's model to show that there are both domain dependant skills and domain independent skills. For example, domain independent skills include communication skills, which are not specific to a particular clinical task, and so may be assessed in a range of situations.

Success at each level of assessment prepares the learner for the next level on their journey to independent clinical practice. Rethans et al[6] argued that performance of a clinician in real life situations is affected by the individual, and by workplace systems, facilities, and resources. 'Does' is a better assessment of a person's performance than 'shows how', and so it is important to assess learners in real-practice, not only in the artificial environment of a university.

One way of evaluating an assessment is to use the utility equation (Figure 2).[7] This is a conceptual model, not a mathematical equation. Reliability considers whether the same result would be achieved if the assessment were repeated. Validity relates to whether we are assessing what

we claim to be assessing. Educational impact can be viewed in terms of what type of learning activity the assessment encourages, and whether developmental feedback is provided to encourage further learning. Acceptability concerns whether stakeholders, such as institutions, healthcare colleagues, learners, and patients, find the assessment acceptable. Finally, the costs, not just in terms of money, but also time, are considered.

If one of these characteristics equals zero, the assessment has no utility. The model also indicates what compromises are made for each assessment, and what elements may need to be improved. For example, an assessment may have good reliability, validity, and educational impact, but it may have poor acceptability, or high costs.

WORKPLACE BASED ASSESSMENTS

Assessment within the workplace and at postgraduate level is less developed than assessment at undergraduate level. The curriculum and learning outcomes are often less well defined, and the context is less predictable, more varied, and more complex. Workplace Based Assessments (WBA) may assess performance of a group of integrated skills, rather than individual skills. They are intended to assess at the upper two levels of Miller's pyramid, 'shows how', and 'does', and often involve clinical performance assessed through observation by more experienced team members.

A WBA may be based on a single encounter or on several encounters. The WBA document designed for the Basic Medical Training (BMT) Logbook, 2016, of the College of Physicians and Surgeons of South Sudan, can be used for both. It provides formative assessment, and evidence towards the assessment of programme completion.

There are several grounds on which workplace performance can be assessed: occurrence, quality, and fitness and suitability.[8] A checklist may be used to note whether a particular behaviour or procedural step has occurred and been observed by the assessor; no judgement is made on quality, just on occurrence. Quality of performance is commonly assessed using a global rating scale. This is what is used in the BMT WBA document, where trainees are rated on a scale from 5 (Well above my expectation of a doctor at current level of training) to 1 (Well below my

Figure 2. The Utility Equation[7]

$$\text{Utility} = \text{Reliability} \times \text{Validity} \times \text{Educational Impact} \times \text{Acceptability} \times \text{Cost}$$

expectation of a doctor at current level of training). Global rating scales can be used whatever the clinical encounter involves. Fitness and suitability assess whether the trainee's performance was satisfactory or unsatisfactory. This can be expanded, as with the BMT WBA, to consider the trainee's level of independent practice.

There are a variety of different formats of WBA that can be found by searching the internet. In the UK, these include the mini-Clinical Evaluation Exercise (mini-CEX), Direct Observation of Procedural Skills (DOPS), and multisource or 360° feedback (MSF). Each considers a different aspect of performance. Trainees are assessed several times over a defined training period, and assessments are collected into a portfolio. Each year the portfolio is used as the basis for an assessment of suitability to progress.

Using a mini-CEX, the trainee is observed in a single patient encounter by a more senior colleague. The trainee's performance is given a score, and feedback is provided. Various elements of competency can be scored, including history taking, physical examination skills, and clinical judgement. DOPS is used in a similar way to assess practical procedures. MSF uses performance over time as the basis of an assessment of general professional skills. It involves the collection of feedback from colleagues including doctors, nurses, and allied healthcare professionals. Eight or more questionnaires are completed anonymously, and the findings collated before being offered to the learner.

ENTRUSTABLE PROFESSIONAL ACTIVITIES

Entrustable Professional Activities (EPA) are tasks that a person with appropriate training and assessment has been entrusted to carry out.[9, 10] An EPA document describes what work is to be carried out and to what standard, for example, to develop and implement a patient management plan. EPAs and competencies can be mapped against each other, to show which competencies are involved in an EPA. EPAs can be used in both undergraduate and postgraduate training. The decision to trust a candidate with undertaking an EPA unsupervised marks mastery of that EPA.[5]

OBJECTIVE STRUCTURED CLINICAL EXAMS

Objective Structured Clinical Examinations (OSCEs)[11] are used extensively within health care education to assess clinical skills. They are designed to assess performance against a standard, in a safe, simulated clinical environment. They are made up of a series of stations, each providing the candidate with appropriate clinical information and equipment to carry out a simulated task. The patient's role is performed by mannequins or actors, or real patients who have volunteered. When used as a summative assessment they assess competence to progress to the next stage of training.

When designing OSCEs, as a first step it is important to decide what will be assessed. 'Blueprinting' is a way of determining the content of an assessment, by checking that the assessment is aligned with the intended learning outcomes, and with the learning and teaching approach. Homer and Russell[12] argue that blueprinting should not just be carried out before OSCE stations have been developed, but also repeated once the stations are written, in order to ensure that constructive alignment has been achieved. If you wish to find out more about blueprinting then Khan et al[13] may be a good place to start.

OSCE stations generally take 5-15 minutes.[13] Ideally all stations should be piloted, and examiners should have assessment training prior to the OSCE. As well as appropriate resources for the clinical simulation, each station requires a set of instructions for candidates, instructions and scoring information for examiners, and instructions and scripts for any patients or role players. Suggestions for how to organise an OSCE can be found in Khan et al.[13]

How many stations are required to ensure reliability? This requires some psychometrics, which are outside of the scope of this article. However, the number of stations will also be affected by the time it takes to complete them all, and how many students can undertake the exam in one day. Having a number of trained examiners, with different examiners in each station, helps reliability.[14] Further information on appropriate psychometrics for OSCEs can be found in Pell et al.[15]

Some institutions undertake a screening OSCE for all students. The results are calculated, and those students on whom you need more information to allow them to progress take a further set of stations. Some programmes also decide that there are a minimum number of stations that a student must pass in addition to the cut score. This prevents students who do well in a few stations but perform poorly in most, from passing the exam. Again, this is outside the scope of this article, but if you are interested read Homer and Russell.[12]

CONCLUSION

In professional training and education, the sum is greater than the individual parts. No one assessment can assess everything a candidate needs to know or demonstrate competency in. Therefore, it is important to look across programmes of learning in order to ensure that the different assessments undertaken, collectively demonstrate that the candidate is competent and safe.

Detailed practical information on how to develop and implement these different types of clinical skill assessments is outside the scope of this article. However, many texts already exist that will help you take the next steps and

we have cited some here. An internet search, prompted by this article, will uncover more. Whatever type of assessment you choose to use, it is important not to forget the general principles and purpose of assessment: what are you assessing, and why?

References

1. Bregazzi R, Bussey S. Teaching and learning in the clinical workplace. South Sudan Medical Journal. 2023;16(1):20-3.

2. Biggs J. Teaching for quality learning at university. Buckingham: SRHE & Open University Press; 1999.

3. Bloom BS. Taxonomy of educational objectives: the classification of educational goals. New York: Longman; 1964.

4. Miller GE. The assessment of clinical skills/competence/performance. Academic Medicine. 1990;65(9):S63-7.

5. Wass V, Archer J. Assessing learners. In: Dornan T, Mann K, Scherpbier A, Spencer J, editors. Medical Education Theory and Practice. Edinburgh: Churchill Livingstone Elsevier; 2011.

6. Rethans J-J, Norcini JJ, Barón-Maldonado M, Blackmore D, Jolly BC, LaDuca T, et al. The relationship between competence and performance: implications for assessing practice performance. Medical Education. 2002;36(10):901-9.

7. Van Der Vleuten CPM. The assessment of professional competence: Developments, research and practical implications. Advances in Health Sciences Education. 1996;1(1):41-67.

8. Norcini JJ, Zaidi Z. Workplace assessment. In: Swanwick T, Forrest K, O'Brien BC, editors. Understanding medical education: evidence, theory, and practice. Hoboken, NJ: Wiley-Blackwell; 2019. p. 319-34.

9. ten Cate O. A primer on entrustable professional activities. Korean J Med Educ. 2018;30(1):1-10.

10. ten Cate O. Entrustability of professional activities and competency-based training. Med Educ. 2005;39(12):1176-7.

11. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. British Medical Journal. 1975;1(5955):447-51.

12. Homer M, Russell J. Conjunctive standards in OSCEs: The why and the how of number of stations passed criteria. Medical Teacher. 2021;43(4):448-55.

13. Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part II: Organisation & Administration, Medical Teacher, 2013; 35(9): e1447-e1463

14. Swanson DB. A measurement framework for performance based tests. In: Hart, IR, editor. Further developments in assessing clinical competence. Montreal: Can-Heal; 1987.p.13 – 45

15. Pell G, Fuller R, Homer M, Roberts T. How to measure the quality of the OSCE: A review of metrics – AMEE guide no. 49, Medical Teacher, 2010;32(10): 802-811