

# Ensuring Clinical Competence in Oman

Anwar Ali Khan

Specialty Training Lead for General Practice, Health Education North-Central and East London, London, UK

## ARTICLE INFO

### Article history:

Received: 20 May 2015

Accepted: 11 June 2015

### Online:

DOI 10.5001/omj.2015.77

Albert Einstein said: “Everything that can be counted does not necessarily count; everything that counts cannot necessarily be counted.” Assessment is arguably the most important method of driving up standards and yet “there is probably more bad practice and ignorance of significant issues in the area of assessment than in any other aspect of higher education”.<sup>1</sup> However, during my 10-year tenure as International Development Advisor (IDA) for the Membership of the Royal College of General Practitioners (International) [MRCGP (INT)] exam in Oman, I have seen a dedicated team grow in confidence in delivering the assessments needed to ensure doctors finishing their residency programs are of a good standard.

The concept of assessment has broadened and has multiple purposes. There has been a move away from comparison/competition among students to what has or has not been learned; a greater emphasis on criterion-referencing than peer-referencing. Table 1 highlights the essential differences between these two models of assessment.

The challenges in criterion-referenced exams include the need to make sure all the judges make the same judgement about the same performances (i.e. inter-rater reliability), constructing appropriate criteria that are evidenced-based and pragmatic, defining criteria unambiguously, applying criteria fairly, and minimizing examiner differences.

There is also a move to a descriptive approach rather than just raw marks, grades, or statistical manipulation. This has led to a greater emphasis on formative assessment to improve practice rather than summative assessment as in the past. Harden<sup>2</sup> uses the bicycle as a useful model when considering the relationship between teaching and evaluation. The front wheel represents teaching and learning while the assessment is represented

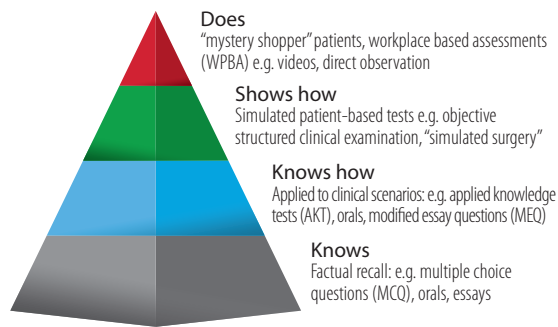
by the rear wheel; problems occur when the wheels go in different directions or are missing. Wass et al,<sup>3</sup> stated that assessment is the most appropriate engine on which to harness the curriculum. Such tests need to be designed addressing key issues such as blueprinting, reliability, validity, and standard setting. Furthermore, there needs to be clarity about whether the assessment is formative or summative. Formative assessment is any assessment for which the priority in its design and practice is to serve the purpose of promoting pupils' learning.

Blueprinting<sup>4</sup> allows the content of the test to be planned against learning objectives since assessment must match the competencies being learned. Columbia University professor of law and education, Jay Heubert, says: “One test does not improve learning any more than a thermometer cures a fever... we should be using these tests to get schools to teach more of what we want students to learn, not as a way to punish them.”<sup>5</sup> Hence, we should use different formats depending on the objectives being tested.

Reliability is the consistency of the test (i.e. how reproducible it is). There are two main facets of reliability: inter-rater and inter-case reliability. Inter-rater reliability is increased by the use of multiple examiners, for example, one question asked by four different examiners is better than one

**Table 1:** Two models of assessment.

Measurement	Standards/mastery
- Norm-referenced	- Criterion-referenced
- Standardized exam	- Determines achievement
- Determines rank order	- Object to separate into two groups (“can do” and “can’t do”)
- Aims to spread out candidates	- Pass rate not predictable
- Usually pre-determined pass rate	- Formative practice easy to do
- Formative feedback limited	



**Figure 1:** The Miller's Pyramid of clinical competence provides an assessment framework.

examiner asking four questions. Inter-case reliability is about candidate performance, and it is known that doctors do not perform consistently from task to task.<sup>6</sup> Consequently, broad sampling across cases is essential to assess clinical competence.

Validity focusses on whether the competencies are tested. Figure 1 shows Miller's pyramid<sup>7,8</sup> of competence that outlines the issues concerning validity. Miller's Pyramid of Assessment provides a framework for assessing clinical competence in medical education and can assist clinical teachers in matching learning outcomes (clinical competencies) with expectations of what the learner should be able to do at any stage. The ultimate goal is to test what the doctor does in the workplace.

There are various methods of setting standards,<sup>5,9</sup> but the choice of method will depend on available resources and consequences of misclassification of passing/failing candidates. MRCGP (INT) Oman has also had to deal with relatively small numbers of candidates when the usual psychometric calculations cannot be relied on fully. They have used a variety of methods to triangulate in order to arrive at a robust pass/fail decision. This is essential since any exam needs to be fair to patients (to ensure the quality of graduates is very good) and to the candidates (by being clear on what is to be tested and how), in that order.

### **Assessment weaknesses**

The multiple choice paper is quite often at the end of the program and one wonders whether this timing could be better. It tests knowledge but not how to use it. We can move up Miller's pyramid by use of applied knowledge (AKT) questions but this move up the pyramid only addresses that students "know

**Table 2:** Deep and surface learning.

Deep	Surface
- Focus is on what is "signified".	- Focus is on the "signs".
- Relates previous knowledge to new knowledge.	- Focus on unrelated parts of the task.
- Relates knowledge from different courses.	- Information for assessment is simply memoried.
- Relates theoretical ideas to everyday experience.	- Facts and concepts are associated unreflectively.
- Relates and distinguishes evidence and argument.	- Principles are not distinguished from examples.
- Organises and structures content into a coherent whole.	- Task is treated as an external imposition.
- Emphasis is internal, from within the student.	- Emphasis is external, from demands of assessment.

how" and do not necessarily extrapolate that the students can apply this knowledge in the workplace.

The oral and modified essay questions (MEQ) allow the candidate to talk or write about things they have learned on courses. But is that what candidates really do? They can "show how" in the simulated surgery, but this can lead to "gaming" by candidates by referring or deferring or even handing out imaginary leaflets to prevent them having to explain their actions.

In the old MRCGP examination, we had a video component and this often produced distortions of real general practice since candidates could select the cases, and there was often artificial "sharing of options". For example, in cases of chest infection, candidates have been known to say "well I can give you amoxicillin or erythromycin, what would you like?" It revealed that candidates were "performing to the test" without understanding why sharing options is good to ensure compliance and concordance with patients.

It has been said that assessment drives learning but I would suggest it has the potential to "skew" learning<sup>10</sup> if candidates are driven to passing exam rather than using them as means to test whether they are providing a good service to patients. Deep and surface are two approaches to learning by candidates [Table 2], this has been derived from original empirical research by Marton and Säljö<sup>11</sup> and was elaborated by Ramsden and Atherton.<sup>12</sup> It is important to clarify what these approaches are not. Although learners may be classified as "deep" or "surface", they are not attributes of individuals:

one person may use both approaches at different times although she or he may have a preference for one or the other. In the UK, many candidates clamor to do exam preparation courses but their best preparation course is their daily practice when they should demonstrate good patient care to themselves and their trainers and, as a side effect, pass the exam with ease.

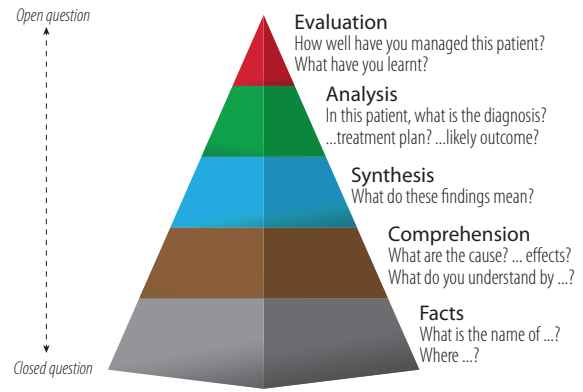
MRCGP (INT) Oman is an excellent exam that has a good reputation internationally of being a robust and challenging, but fair. However, I would suggest that we improve assessment of trainees by reducing the burden of assessment, thinking about the timing of the assessments, and using tools that are better aligned with the curriculum.

Workplace-based assessments (WPBA) provide such an opportunity to recouple teaching and learning with the assessment. Its authenticity is good as it is as close as possible to the real situation in which the trainees work. Furthermore, it provides us with the only route into many aspects of professionalism; trying to test this in a simulated surgery, for example, only shows “how to pass the station” not what they really do.

WPBA allows the integration of assessment and teaching and uses tools that include feedback. They are, however, summative in that they count but yet they are very formative. It allows both the trainer and learner to be much clearer where they are, not only about their learning trajectory but also their learning needs. It involves an honesty in the relationship but importantly, assessments are no longer “terminal” but can allow reflection by the trainee and hence further their development. Such assessment becomes formative assessment when the evidence is used to adapt the teaching work to meet learning needs.<sup>13</sup>

This does mean having an educator workforce conversant and confident in the use of methods that move away from the didactic course of teaching towards a Socratic method and heuristic style, which will then promote the development of the reflective practitioner by promoting self-assessment.<sup>14</sup> Figure 2 demonstrates the use of appropriate questioning to “get under the skin” of the learner and thereby help his/her progression.<sup>15</sup>

A detailed discussion of the various methods in WPBA is beyond the remit of this article, but I hope it will stimulate greater interest in this important aspect of educating and assessing the future doctors in Oman.



**Figure 2:** Hierarchy of knowledge with examples of questions to determine the learner’s knowledge.

In summary, the development of robust WPBA for the residency programs will bode well for a competent workforce delivering the best possible, cost-effective care that patients in Oman deserve. However, this does need a strategic view and deployment of appropriate resources to develop the required educator workforce to put this vision into practice. Furthermore, there needs to be a change in attitudes amongst some teachers and most learners that promotes an atmosphere of inquisitive questioning rather than simple blind obedience to the teacher. Finally, I would remind all you trainers and teachers of the Hebrew proverb: “Do not confine ‘students’ to your own learning for they were born in another time.”

#### Disclosure

No conflicts of interest, financial or otherwise, were declared by the author.

#### REFERENCES

1. Boud D. Assessment and learning: contradictory or complementary. In: Knight P (Ed). *Assessment for Learning in Higher Education*. London: Kogan. 1995. p. 35-48.
2. Harden RM. Assessment, Feedback and Learning. In: *Approaches to the Assessment of Clinical Competence*. International Conference Proceedings. Norwich: Page Brothers. 1992. p. 9-15.
3. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet* 2001 Mar;357(9260):945-949.
4. Dauphinee D. Determining the content of certification examinations. In: Newble D, Jolly B, Wakeford R, eds. *The certification and recertification of doctors: issues in the assessment of clinical competence*. Cambridge: Cambridge University Press, 1994. p. 92-104.
5. Boston Globe. Testing expert calls for more than MCAS results to assess students. Available at: <http://www.fairtest.org/organizations-and-experts-opposed-high-stakes-test>. Accessed December 7, 2001.
6. Swanson DB, Norman GR, Linn RL. Performance-based

- assessment: lessons learnt from the health professions. *Educ Res* 1995;24:5-11.
7. Norcini JJ. Work based assessment. *BMJ* 2003 Apr;326(7392):753-755.
  8. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990 Sep;65(9)(Suppl):S63-S67.
  9. Cusimano MD. Standard setting in medical education. *Acad Med* 1996 Oct;71(10)(Suppl):S112-S120.
  10. Khan A, Jamieson A. Introduction. In: Jackson N, Jamieson A, Khan A, editors. *Assessment in Medical Education and Training – a practical guide*. New York: Radcliffe publishing; 2007.
  11. Marton F, Saljo R. On qualitative differences in learning – 1: outcome and process. *Br J Educ Psychol* 1976;46:4-11.
  12. Atherton J. Learning and Teaching: Deep and Surface learning. Available at: <http://www.learningandteaching.info/learning/deepsurf.htm>. Accessed 2005.
  13. Black P, Harrison C, Lee C, Marshall B, William D. *Assessment for Learning: Putting it into Practice*, Maidenhead. Open University Press; 2003.
  14. Cioffi J. Heuristics, servants to intuition, in clinical decision-making. *J Adv Nurs* 1997 Jul;26(1):203-208.
  15. Lake FR, Vickery AW, Ryan G. Teaching on the run tips 7: Effective use of questions. *Med J Aust* 2005 Feb;182(3):126-127.