OSCEs are outdated: clinical skills assessment should be centred around workplace-based assessments (WPBAS) to put the ‘art’ back into medicine [version 1]

Hamed Khan
St Georges

Abstract
This article was migrated. The article was marked as recommended. OSCEs have gradually replaced 'long cases' as the mainstay of undergraduate clinical skills assessment because of their objectivity, consistency and reliability. But the aspects of OSCEs which make them so reliable increasingly encourage students to prepare strategically, who often adopt a robotic 'tickbox' approach, rather than use OSCEs as a tool to learn clinical skills for safe competent real-life practice. Thus, whilst OSCEs facilitate technical competence, they do not prepare students for the unique nuances that make medicine an 'art' as well as a science. In pursuit of consistency and reliability, we are sacrificing validity and not preparing future doctors for the innate nuances and variability that make medicine so unique- and which often come as a shock to newly qualified doctors who orientate their undergraduate learning around OSCEs rather than real life. The doctors of the future will need to be adaptable and be able to vary their practice depending on the clinical and biopsychosocial context much more so than before. To drive their learning accordingly, we need a paradigm shift in medical education and assessment. WPBAs should now take centre-stage in undergraduate clinical assessment, with OSCEs significantly scaled back.

Keywords
OSCE, Assessment, Clinical Skills, Workplace based assessments, Medical Education

Open Peer Review

Migrated Content
"Migrated Content" refers to articles submitted to and published in the publication before moving to the current platform. These articles are static and cannot be updated.

1. Wolf Hautz, Inselspital Bern
2. Julie Browne, Cardiff University School of Medicine
3. Trevor Gibbs, AMEE
4. Megan Anakin, University of Otago
5. Paul Duggan, The University of Adelaide
6. P Ravi Shankar, American International Medical University

Please see the article web page for additional reviewers.

Any reports and responses or comments on the article can be found at the end of the article.
OSCEs are outdated: clinical skills assessment should be centred around workplace-based assessments (WPBAs) to put the ‘art’ back into medicine

I love OSCEs. I have been fascinated by them since I started medical school. As a student, I voluntarily spent several hours of my spare time setting up mock OSCEs for fellow students, when I should have been studying myself. After I qualified, I became an OSCE examiner for 3 different London medical schools. Later in my career, I wrote an OSCE textbook. Now, at the pinnacle of my career as a Lecturer in Clinical Skills, I am in charge of running an OSCE for two entire cohorts of students at a prominent London medical school, doing what was always my ‘dream job’.

So I have been immersed in OSCEs all my life- and have thoroughly enjoyed it. Setting up an OSCE, watching its mechanics roll into action as intelligent but terrified students demonstrate their skills and knowledge, and then seeing that translate into a structured report on their performance, still gives me unparalleled joy.

However over the last few years, my love affair with OSCEs has soured. I have become increasingly sceptical about how well OSCEs prepare students for the trials and tribulations of real life practice. I reluctantly concluded that OSCEs were no longer fit for purpose when I met one of former students who had just graduated. He was a brilliant, hard working, insightful student, so I was surprised when he told me that he was terrified of starting work as a doctor. When I asked him why, he gave me an interesting answer- he said that all of his life, he had been learning clinical skills to pass OSCEs- and now, he needed to learn clinical skills for real life.

Firstly, one cannot underestimate the decisively positive impact that OSCEs have had on clinical learning. Decades ago clinical skills assessment was based around ‘long cases’, with relatively open-ended instructions, examiners who could ask questions and conduct vivas as they wished to. The inevitable examiner bias led to significant unreliability.

I imagine that this unreliability drove the evolution from ‘long cases’ to the modern day OSCE, with specific tasks and instructions for candidates, and clear objective marking criteria for examiners. Both candidates and examiners are briefed carefully, and academics work relentlessly to ensure consistency and reliability.

But in my opinion, the pendulum has swung too far. As academics, our well intentioned focus on process has led us to neglect crucial outcomes, and we are losing the forest for the trees.

Our focus on reliability, and making OSCEs as objective and structured as possible, has led to a paradigm where students know what to expect. There is little variation amongst OSCE scenarios both within and across medical schools, and OSCE textbooks and resources contain very similar content. Unsurprisingly, students largely orientate their learning strategically to pass OSCEs, often adopting a robotic ‘tickbox’ approach. The idea of using OSCEs as a tool for meaningful learning to become safe competent practitioners who are able to assess and treat real patients rarely motivates students- as demonstrated by my example above.

Hence our students gain technical competence from OSCEs, but struggle to translate this into skills which are practically useful in the dynamic, ever changing environment of real frontline medicine. The A&E department with several acutely sick patients presenting simultaneously, or the inner city GP practice where doctors deliver holistic care for patients whilst disentangling medical issues from psychosocial factors, is far removed from the comfortable confines of an OSCE station, where simulated patients always stick to the script, and candidates never really need to expect the unexpected. The gap between the ‘Shows How’ of Miller’s Pyramid, and ‘Does’ in clinical practice is getting increasingly steeper. In real life, a plethora of variables intertwine with pathology to create the unique dynamic of individual scenarios and consultations- such as psychosocial factors, patient perspectives, time and resource pressures, in addition to natural variation amongst individual humans. Whilst OSCEs help students attain basic competence in clinical skills, they fail to prepare them for the complex nuances that make medicine unique amongst disciplines, as both an ‘art’ as well as a science. If anything, these nuances will only become more complex for the doctors of the future, as they face unique challenges- aging populations with complex medical problems intertwined with intractable social issues, resource constraints, and an unprecedented increase in access to information, leading to patients who are better informed, and treatment guidelines and pathways which change on a daily basis.

Thus whilst technical competence will remain crucial, the doctors of the future will also need to be able to adapt and improvise their skills depending on the situation, much more so than ever before. They will need to be able to innovate and think outside the box, both on the clinical front line, as well as outside it as leaders and managers of their health services.
In my opinion, whilst we have made vital progress with assessment reliability from OSCEs, we are sacrificing validity to an unacceptable extent, and leaving our future doctors unprepared for the innate nuances and variability that make medicine unique.

So what is the answer? A key overarching principle of high quality assessment is having a variety of both assessments and assessors. But more specifically, I feel that a paradigm shift is needed where we put workplace-based assessment (WPBA) at the centre of clinical skills assessment, and significantly reduce the role of OSCEs. WPBAs such as CEXs and CBAs are used extensively for postgraduate assessment. They allow us to assess both technical competence, as well as the ability of students to adapt skills, and improvise and innovate, and vary their practice depending on the individual dynamic and context of the scenario. Using WPBAs also enforces a vital ‘hidden curriculum’ message that the ultimate aim of learning clinical skills is for real life practice with real patients, rather than confining it to artificial simulated settings, and thus change the way students think- both individually and collectively. Students will feel compelled to think more about how they would realistically conduct a cardio-respiratory exam in an elderly lady who is short of breath in A&E, and less about the causes of clubbing and janeway lesions- which are rarely seen in practice, but seem to be crucially important for OSCEs. Over the course of time, those with a more collateral role in student education, such as textbook authors, private course organisers, societies etc, will also change their approach to students, and a paradigm shift will slowly occur.

Whilst I am advocating a reduced role for OSCEs, I am absolutely not suggesting that they be discarded altogether. OSCEs play a vital role in assessing basic technical competences, and the reliability that they bring to medical assessment overall is also crucially important.

I also recognise the drawbacks of WPBAs, such as interassessor variability, and the logistical complexities of conducting assessments on a large scale in ‘live’ clinical settings- but we have a precedent. WPBAs were introduced, and are now used extensively in postgraduate training and assessment in the UK. The Royal College of General Practitioners, which sets the assessments used to certify the completion of GP training, uses WPBAs extensively, and GP training is now unimaginable without them. This is particularly noteworthy as most care in the British NHS is delivered through GPs, which is expected to be the end career point for most medical students. Surely we can look at what similar postgraduate training organisations have done, and use, adapt and improve on their work and their existing infrastructures, to build similar frameworks for undergraduate clinical assessment?

At the very least, we need to start dedicating time, thought and resources towards developing frameworks and educators with the aim of making WPBAs much more integral to clinical skills than they are now. The debate around clinical skills assessment has stagnated, and the perception that OSCEs are the panacea of clinical skills assessment is increasingly unquestioned. We need to reopen that discussion, and start pushing its boundaries in order to reignite the evolution which led to the formation of OSCEs in the first place. It is high time that we used WPBAs to compel our students to step outside the comfort zones of surreal OSCE settings, and learn and practice their clinical skills to excel in real life settings, with real patients, in real clinical scenarios.

**Take Home Messages**

- OSCEs are outdated, and encourage strategic OSCE orientated learning, rather than ‘deep’ meaningful learning to become safe competent clinicians.

- In our well-intentioned pursuit for reliability, we have made huge sacrifices in validity. OSCEs no longer prepare students for the innate nuances that make medicine so unique as both an ‘art’ as well a ‘science’.

- Factors such as aging populations, financial constraints, and ever improving access to information (for both patients and doctors) mean that future doctors will need ‘real life’ skills such as the ability to adapt, improvise and innovate more so than ever before- both as service providers as well as managers and leaders of their health services.

- Workplace based assessment allow us to assess students ability to do this as they are centred around real life patients in real life settings, rather than in the comfort zone of simulated OSCE settings.

- It is time for academics in undergraduate medicine to take a leap of faith, and work towards making WPBAs the mainstay of clinical skills assessment, and scale back the role of OSCEs.
Notes On Contributors
Dr Hamed Khan is a Lecturer in Clinical Skills and Responsible Examiner for the Clinical Sciences OSCE at St Georges, University of London. He has been an OSCE examiner at three London medical schools, and has written a textbook on OSCEs for medical finals (http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0470659416.html). In addition to this, he is a practising GP in South London, and is a National Council member of the British Medical Association. He has also written on NHS related issues for the Guardian and the Channel 4 websites.

Declarations
The author has declared the conflicts of interest below.
I am a Lecturer in Clinical Skills at St Georges (University of London) and the Responsible Examiner for the Clinical Sciences OSCE. I am also a GP in London. I have co-authored a textbook on OSCEs for Medical Finals. I presented a ‘Point of View’ oral presentation at the AMEE 2017 Conference.

Acknowledgments
PLEASE NOTE: this article and its abstract are derived from my ‘Point of View’ abstract and presentation at the AMEE 2017 Conference in Helsinki.
Open Peer Review

Migrated Content

**Version 1**

Reviewer Report 29 September 2018

https://doi.org/10.21956/mep.19377.r28379

© 2018 David.M A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Anthony David.M**

Soho

This review has been migrated. The reviewer awarded 4 stars out of 5

I enjoyed reading the opinion piece and the interesting reviews too. Indeed the author is passionate and presents his case very well. It is a fact that the Global assessment is sadly lacking in OSCEs. The OSCEs have come in as a reaction to the subjective and unreliable methods of clinical assessment which existed in the form of long and short cases. But it has its own pitfalls. Workplace based assessment methods seems to be the more suitable method. But as some earlier reviewers opined, an eclectic approach using more than one method of assessment may be best.

**Competing Interests:** No conflicts of interest were disclosed.

Reviewer Report 07 April 2018

https://doi.org/10.21956/mep.19377.r28385

© 2018 Gunaratne N. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Nandalal Gunaratne**

Faculty of Medicine, Wayamba University of Sri Lanka

This review has been migrated. The reviewer awarded 3 stars out of 5

The Title of the article is quite striking but the article itself shows a more balanced opinion from the author. Perhaps we could lend more time to improving and expanding the OSCE.WPBA has its own
pitfalls and is not as suitable for assessing undergraduates as it is in assessing the postgraduate. This is because the undergraduate is not a "real doctor" and has a different relationship with both patients and staff.

**Competing Interests:** No conflicts of interest were disclosed.

Reviewer Report 05 November 2017

https://doi.org/10.21956/mep.19377.r28381

© 2017 Jaap A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Alan Jaap
University of Edinburgh Medical School

This review has been migrated. The reviewer awarded 2 stars out of 5

Thank you for this thought-provoking commentary which has unsurprisingly stimulated a healthy debate on the difficult area of how to assess clinical performance authentically while making best use of the limited resources available. The author highlights the problems associated with relatively simple OSCE stations and check list marking. However, many would argue that the educational impact of such an approach is actually quite positive for early clinical learners, and, as other reviewers have already stated, the complexities of both station design (e.g. integrated skills) and marking schedules (e.g. domain based) can be increased to make the assessment more authentic as students progress towards graduation. The problems of existing WPBAs have also been outlined by several other reviewers, and I would simply add that it remains difficult to ensure consistent sampling of learning outcomes in the workplace, as by their nature these WPBAs are used opportunistically. A move to more longitudinal assessments in the workplace, for instance through the introduction of a framework of Entrustable Professional Activities, may offer a feasible solution to many of the issues raised going forward.

**Competing Interests:** No conflicts of interest were disclosed.

Reviewer Report 03 November 2017

https://doi.org/10.21956/mep.19377.r28388

© 2017 Flaherty G. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Gerard Flaherty  
National University of Ireland Galway  

This review has been migrated. The reviewer awarded 3 stars out of 5  

The author provides a timely commentary on the future role of workplace-based assessment in undergraduate clinical skills education. His credibility stems from his vast experience in OSCE station design and in OSCE coordination and delivery. While I agree with his sentiments about the strategic learning which OSCEs can inspire, I wonder if he is being unfair on students, whose motivation probably extends well beyond merely preparing for OSCEs. Our challenge in clinical skills assessment is to maximise the value of OSCEs so as to encourage deep learning, and to remove the predictability which the author criticises. In this respect, it behoves us to collaborate in designing more authentic stations, perhaps of longer duration, and in utilising simulation-based assessment to a greater extent. This would enable us to assess with greater sophistication students' human factors skills, including situational awareness and multi-tasking. This is happening to varying extents already in different medical schools, but I believe there is untapped potential in sharing best OSCE design practice. Perhaps the model has become stale and has lost its creativity. OSCEs remain a very useful, efficient and reliable tool for undergraduate clinical skills assessment, but we need to do more to adapt them to changing healthcare needs, by designing more complex, authentic and challenging stations. This opinion paper is thought-provoking and may be a stimulus for more original research in collaborative OSCE station design, and I therefore welcome its submission.

**Competing Interests:** No conflicts of interest were disclosed.

Reviewer Report 02 November 2017  
https://doi.org/10.21956/mep.19377.r28377

© 2017 Pinnock R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ralph Pinnock  
University of Otago, Dunedin, New Zealand  

This review has been migrated. The reviewer awarded 3 stars out of 5  

I have a lot of sympathy with the views expressed by Dr Khan but wonder whether the answer is not ‘yes’ to both forms of assessment as components of programmatic assessment. OSCEs are not easy to organise in children. Although there publications of using children as simulated patients in OSCEs to assess medical students’ clinical skills, they can be difficult to organise. How many candidates can a well 5 year old child be expected to see? When I became responsible for a paediatric curriculum where I
previously worked we changed the in course assessment from an end of rotation OSCE to 5 miniCEXs during the rotation. We used patients on our inpatient service because we wanted to assess our students’ clinical skills with children who had abnormal physical signs. Our students were often reluctant to see children who had been admitted acutely citing the stress that both mother and child would be experiencing. To try and allay students concerns we interviewed mothers and children to find out their attitude to seeing medical students (Pinnock R, Weller J, Shulruf B, Reed P, Mizutani S. Why parents and children agree to become involved in medical student teaching. (J. Paediatr.Child Health 2011;47(4):204 - 210). I was surprised how even children as young as 6 or 7 were willing to see and be examined by students so that “doctors can learn to help other children”. It was always necessary to see the parent and child after the miniCEX because almost invariably the parent had become concerned by some of the students comments usually, related to including a very serious disorder in the differential diagnosis. One very valuable insight gained from this is work-placed base assessment into the student's performance was to ask the parent and child if they would see that “doctor” again. There are certainly advantages to work-based assessments.

**Competing Interests:** No conflicts of interest were disclosed.

[Reviewer Report 01 November 2017](https://doi.org/10.21956/mep.19377.r28380)

© 2017 Hunt J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Julie Hunt
Lincoln Memorial University

This review has been migrated. The reviewer awarded 2 stars out of 5

As a clinical skills educator who creates OSCE materials, I read this opinion piece with interest. The author proposes that OSCEs be eschewed in favor of WPBA which he suggests have higher validity albeit lower reliability. As previous reviewers have stated, I would have liked to have seen references in support of some of his points. Different forms of assessment have varying strengths and weaknesses in terms of feasibility, reliability, validity, richness of feedback, and suitability for different levels of learners. Assessments of multiple different types probably give us the most complete view of a student's competence, as has been proposed by van der Vleuten.

**Competing Interests:** No conflicts of interest were disclosed.
Richard Hays  
James Cook University

This review has been migrated. The reviewer awarded 2 stars out of 5

While I found this opinion paper entertaining and thought provoking, I wonder the author has been distracted by popularism rather than theory. OSCEs play an important role in demonstrating that, at a particular point in time, candidates demonstrate the ability to perform tasks that are essential components of future practice. It is possible to go beyond this point, with more elaborate simulations and tasks. Workplace based assessment is really one of those ‘holy grail’ issues. Yes, what candidates actually do in the workplace in ultimately more important, and should be part of the assessment package, but there are many pitfalls. The first is that scores have been shown to be inflated, with very similar rank ordering, and so standard setting is more difficult. Most achieve what appear to be high scores, although norm referencing can be helpful. Second, working with colleagues whom we assess raises the likelihood that judgements are based on global, halo impressions of being ‘nice’ and ‘good’, or possibly the opposite. This is where validity and reliability have the potential to collide. The third problem is the answer to the last point. WBA requires multiple judges, multiple assessments, multiple methods on multiple occasions and from different perspectives. This is just really hard to do well. If we are not careful, we will replace OSCEs with narrowly based judgements by colleagues, which would be almost full circle in the development of assessments in medical education. All very thought provoking, but there is unlikely to be any simple solution.

**Competing Interests:** No conflicts of interest were disclosed.

---

P Ravi Shankar  
American International Medical University

This review has been migrated. The reviewer awarded 4 stars out of 5

I read with great interest this thought provoking paper. When I was in medical school, long cases and
short cases were the mainstay of clinical exams. Long cases may more accurately reflect the variability and the nuances of medical practice but the major problem was variability in grading among examiners, differences among patients and among diseases. The trend toward OSCE came about primarily because of the need for objectivity and reliability. For high stakes exams it became important to reduce bias and variation. The premise was the process of medical consultation could be broken down into various component parts which could be examined separately and objectively. However, many authors now think that this may be an oversimplification. Caribbean medical schools mainly prepare students for a career in the United States and Canada and OSCEs and standardized patients are important methods of assessment. Standardizing variables in the patient-doctor consultation makes assessment more standardized and reliable but as the author mentions the variability and the 'richness' of a real-life consultation may be missing. WPBA may be a good alternative and has been widely used in postgraduate education. This may however, need to be tailored at a different level during undergraduate education as the expected skill sets and/or the required level of these skills may be different. I agree with other reviewers that adding references will further strengthen the manuscript.

**Competing Interests:** No conflicts of interest were disclosed.
**Competing Interests:** No conflicts of interest were disclosed.

**Reviewer Report 23 October 2017**

https://doi.org/10.21956/mep.19377.r28376

© 2017 Browne J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Julie Browne**
Cardiff University School of Medicine

This review has been migrated. The reviewer awarded 4 stars out of 5

I thoroughly enjoyed this article and found myself almost wanting to jump to my feet and cheer at its rousing call for ‘real life’ educational experiences and assessment practices that accept the variability of genuine clinical experience rather than attempting to homogenise it. I am sure many other medical educators who have struggled for some time to encourage their students to get out of the simulation suite and onto the wards will agree with its principal premise: that if you examine students through simulation then that’s how they will prefer to learn, because “assessment drives learning”. Furthermore, I was pleased that its argument was rational and proportionate: it outlined the many unintended consequences of an over reliance on OSCEs while freely admitting the enormous challenges involved in reducing the supremacy of OSCEs as the default assessment method, especially for large scale, high stakes tests of clinical skills. I am looking forward to seeing educators rising to the challenge it poses – namely, how to develop real life assessments that can offer sufficient validity and reliability to ensure both student learning and patient safety. Just a final point - I wonder why this article has no references? It would be strengthened considerably if it were properly situated within the existing debates around simulation, workplace based assessment and education for patient safety. I would, for example, have expected to see reference to one of the earliest papers that kicked off this whole debate: Bligh J, Bleakley A. Distributing menus to hungry learners: can learning by simulation become simulation of learning? Med Teach. 2006 Nov;28(7):606-13; but there have been many more in the 12 years since it was first published.

**Competing Interests:** No conflicts of interest were disclosed.

**Reviewer Report 23 October 2017**

https://doi.org/10.21956/mep.19377.r28382

© 2017 Anakin M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Megan Anakin  
University of Otago

This review has been migrated. The reviewer awarded 2 stars out of 5

I was interested in reading this article because I come from an education background and I, too, have experienced how powerful assessment can be as a driver of student learning. In addition to the suggestions made by the previous reviewers, a discussion of how the different principles of assessment are at play – and potentially creating the set of tensions that the author and student experienced would enhance article. The vivid anecdote about the student and persuasive argument could be augmented by referencing the literature examining students’ experiences of assessment generally, and summative medical assessments, in particular. The author may find that there is a paucity of medical education literature in this area and the need to investigate this issue further may strengthen his argument.

**Competing Interests:** No conflicts of interest were disclosed.

Reviewer Report 23 October 2017

https://doi.org/10.21956/mep.19377.r28384

© 2017 Gibbs T. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Trevor Gibbs  
AMEE

This review has been migrated. The reviewer awarded 2 stars out of 5

As someone too who has grown up with OSCEs and can remember the time before OSCEs, I was drawn to read and review this paper. Although I agree that we have to be very careful how we use OSCEs and at what point in time in students’ life, to actually make a statement that OSCEs are outdated and encourage strategic OSCE thinking is as far as I can tell from research completely unfounded. There is nothing wrong with a correctly structured OSCE, performed and organised in a correct manner and used to specifically assess students at a specific point in their career. OSCE's do not give you a real picture of how the students perform but they do have good predictive validity Graham et. al J Dent Educ. 2013 Feb;77(2):161-7)My argument is not the value of the OSCE but the way in which it is used or misused, and specifically how it fits into the toolbox of assessment, and how it is used with other forms of assessment to give as near as possible realistic picture of the student. I would have preferred to have seen some more evidence supporting this very strong statement

**Competing Interests:** No conflicts of interest were disclosed.
This is an interesting viewpoint, that elaborately argues why OSCEs should be replaced with workplace based assessments. The authors key reason is, that - in his opinion - the OSCEs lack clinical validity and foster learning the wrong thing. Both arguments are worth consideration in my opinion. However, in advocating the replacement of OSCEs with WPBAs, the author neglects the very different use these assessment formats have. OSCEs are most often used for summative purposes and are thus required to be highly reliable (which, depending on their design and the number of stations used, they often are). WPBAs are mostly - and I would add hopefully - used for formative assessment, i.e. providing feedback. As such, they should be valid, while reliability is less of a concern. Thus, OSCEs are mainly an assessment tool, while WPBAs are basically teaching tools. Replacing one with the other would mean to use the wrong tool for the task.

*Competing Interests:* No conflicts of interest were disclosed.