



Harnessing insights from an activity system – OSCEs past and present expanding future assessments

Helen Reid , Gerard J. Gormley , Tim Dornan & Jennifer L. Johnston

To cite this article: Helen Reid , Gerard J. Gormley , Tim Dornan & Jennifer L. Johnston (2020): Harnessing insights from an activity system – OSCEs past and present expanding future assessments, Medical Teacher, DOI: [10.1080/0142159X.2020.1795100](https://doi.org/10.1080/0142159X.2020.1795100)

To link to this article: <https://doi.org/10.1080/0142159X.2020.1795100>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 31 Jul 2020.



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



Article views: 620



View related articles [↗](#)







View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Harnessing insights from an activity system – OSCEs past and present expanding future assessments

Helen Reid^a , Gerard J. Gormley^a , Tim Dorman^{a,b}  and Jennifer L. Johnston^a 

^aCentre for Medical Education, Queen's University Belfast, Belfast, UK; ^bDepartment of Education Development and Research, Maastricht University, Maastricht, The Netherlands

ABSTRACT

Objective Structured Clinical Examinations (OSCEs) are a dominant, yet problematic, assessment tool across health professions education (HPE). OSCEs' standardised approach aligns with regulatory accountability, allowing learners to exchange exam success for the right to practice. We offer a sociohistorical account of OSCEs' development to support an interpretation of present assessment practices. OSCEs create tensions. Preparing for OSCE success diverts students away from the complexity of authentic clinical environments. Students will not qualify and will, therefore, be of no use to patients without getting marks providing evidence of competence. Performing in a formulaic and often non patient-centred way is the price to pay for a qualification. Acknowledging the stultifying effect of standardising human behaviour for OSCEs opens up possibilities to release latent energy for change in medical education. In this imagined future, the overall object of education is refocused on patient care.

KEYWORDS

OSCE; assessment; standardised patients; clinical; undergraduate

Introduction

"They don't say stuff like the SPs [simulated patients] do."

The student quoted here was giving her reason for avoiding clinical contact in the run-up to a final Objective Structured Clinical Examination (OSCE), passing which would allow her to practise as a doctor. She found real patients confusing because they are not standard enough. To us as clinicians, who regard responding to non-standard situations as the essence of our practice, this is a 'call to arms' to question the dominance of OSCEs in health professions education (HPE). Assessment is, of course, inevitable. Being able to pass difficult assessments distinguishes education for the professions from training for occupations (Freidson 1970). Assessments define professional roles, as distinct from the more basic capabilities needed to perform manual occupations. The health professions have led the way in researching and developing the use of assessments to regulate entry, progression, and certification.

Selecting the best form of assessment matters, not just to learners who are keen to progress. It matters to teachers and, particularly nowadays, to regulators, because success in assessments is a key to the door to professional recognition and status. Regulators have to use transparent processes to control entry, registration, and progression (Goodwin 2018) in order to be politically accountable. The United Kingdom regulator, the General Medical Council (GMC), for example, has responded to this pressure by introducing a national licensing examination (NLE) with effect from 2023.

It is in this political climate that OSCEs have been so successful. 'New public management' (NPM) refers to the

Practice points

- OSCEs are widespread across health professions education (HPE).
- OSCEs are a practical manifestation of a paradigm shift towards standardisation and reliability.
- OSCEs can create tensions as learners strive to demonstrate behaviours in pursuit of marks.
- OSCEs risk diverting learners away from non-standard, authentic patient encounters in real clinical environments.
- Harnessing tensions around OSCEs offers potential to refocus assessment on patient care.

adoption of market principles into public services to limit cost and the use of defensible procedures for purposes of political accountability. In the language of activity theory (AT), which this article invokes to critique the contemporary use of OSCEs, NPM shifts attention from the use value (literally and the usefulness) of education to its exchange value (the ability to trade examination performance for capital). Under NPM, OSCEs turn performed behaviour into statistical capital that students can exchange for the social capital of practising medicine. Regulators exchange transparent and defensible assessment procedures for capital, which empowers them to regulate professions.

From the critical position taken in preceding paragraphs, it is logical to use AT to critique OSCEs' position within HPE. AT, as authors of companion articles in this issue illustrate, focuses on human agency within social structures

CONTACT Tim Dorman  t.dorman@qub.ac.uk  Centre for Medical Education, Whitla Medical Building, Queen's University Belfast, Belfast BT9 7BL, UK

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

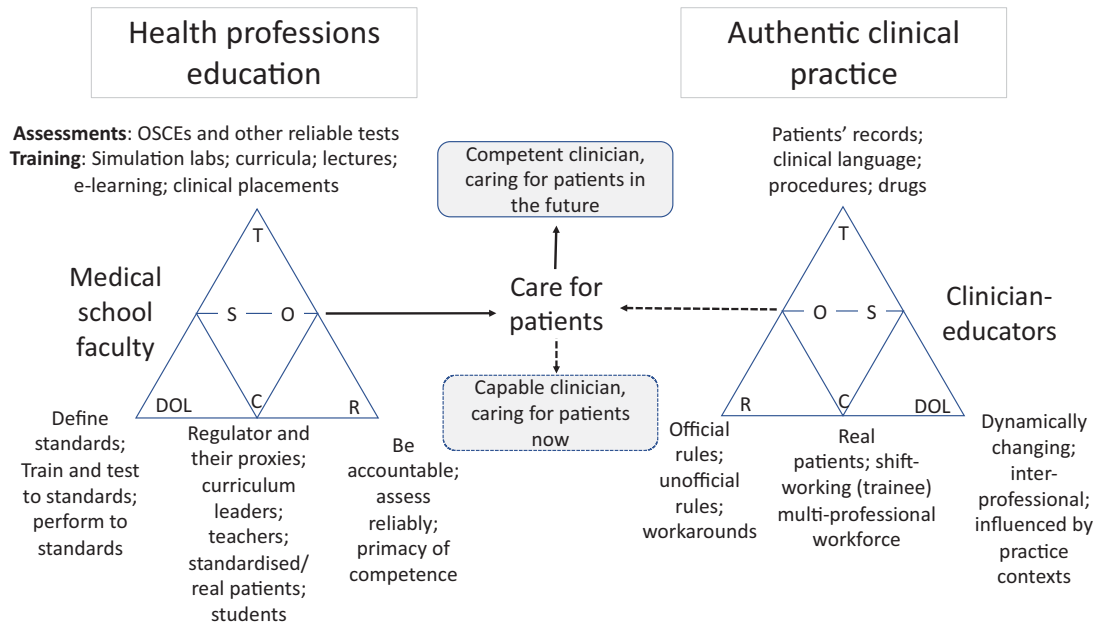


Figure 1. OSCEs as a tool within an activity system of HPE.

and processes. AT offers a dynamic way of examining OSCE practices, in light of their social, cultural, and historical origins. With roots in Soviet dialectical learning theory, AT is optimistic and forward-looking because it views contradictions and tensions as drivers for change.

First, we give an overview of OSCEs' historical development. Second, we use the lens of AT to interpret empirical data representing contemporary OSCE practice. Third, we look to the future and consider how AT can help us expand the system and open new frontiers in HPE, including assessment.

Past: Where did OSCEs come from and what have they stood for in the field of HPE?

Assessment in HPE has taken many different forms in the time since it first took place in Paris in 1788 (Lesky 1970). At first, recall of medical knowledge in written examinations prevailed. Later, tests of clinical performance such as 'long cases' (Ponnampereuma et al. 2009) came to accompany written knowledge tests. Against that background, education developers in a Scottish medical school in the 1970s devised OSCEs, which quickly entered the mainstream of assessment (Harden et al. 1975; Harden and Gleeson 1979). While various written tests remain widespread, the influence of OSCEs quickly turned to dominance. Examiners gave them centre stage in medical school finals and many exit assessments from postgraduate training embraced OSCEs. Since assessment drives learning, clinical education, also, became OSCE-oriented. Researchers, meanwhile, turned assessment in general, and OSCEs in particular, into one of the most thoroughly explored areas of HPE scholarship (Rotgans 2012).

The rise to dominance of OSCEs has a historical as well as a political explanation. The structuring of a broad, predefined range of subject matter offered a solution to something that was plaguing assessment: so-called 'content specificity'. Long cases were unreliable because the same candidate would be expected to perform differently in different cases and with different examiners. Ensuring that all candidates were exposed to standardised, and suitably

broad, content and many different examiners increased the reliability of practical testing, just as multiple-choice questionnaires had proved a more reliable way of testing knowledge than essays or viva voce examinations. OSCEs gained traction and spread across professions, stages of education, and geographical settings to become the globally dominant assessment modality they are today. OSCEs had global impact because they allowed educators, for the first time, to assess practical performance reliably at a time when reliability was sorely needed.

Over time, OSCEs diversified into many related forms of assessment, each with its own acronym: for example, Objective Structured Practical Examinations (OSPES) and Objective Structured Long Examination Records (OSLERs). Selection for admission to medical school adopted Multiple Mini Interviews (MMIs), which have been characterised as 'Admissions OSCEs' (Eva et al. 2004). This standardising and structuring of assessments across multiple 'stations', united by the goal of making decisions about entry to and progression through careers in the health professions, led us to critique 'OSCEs' (rather than 'the OSCE'). It is a whole movement over a period of recent history rather than one specific assessment technology that we discuss here, albeit using an example of a classical OSCE as the topic of empirical research.

Present: OSCEs as a tool in an activity system in HPE

Conceptual lens

The preceding socio-historical review explains how OSCEs became a key tool in activity systems, first of medical education, then of HPE more widely (Engeström and Pyörälä 2020). Figure 1, which we now explain, places the activity system of HPE alongside the activity system of authentic clinical practice.

OSCEs, along with other assessments, curricula, and learning resources, are tools. The subjects are faculty, whose object is equipping students to provide patient care. The dominant (i.e. most powerful) figure in the community of HPE is the regulator. Other members of the

community include curriculum leaders, students, and SPs. Labour is divided so that regulators define standards of competence against which medical schools can assess students to satisfy politicians that they will be fit to provide patient care. Curriculum leaders implement OSCEs in their own curricula. The role of SPs is to be surrogates for real patients. The activities of the community are determined by rules of accountability and the elimination of unsafe practice. The most influential rule is that assessments must be reliable enough to support regulatory processes and, ultimately, lawyers defending a medical school that has deprived a student of their future livelihood. Reliability is assured by standardising subject matter and procedures. The outcome of the activity system is statistically defensible evidence of competence, whose exchange value is students being allowed to start caring for patients and politicians being satisfied with regulators' performance.

The activity system of authentic clinical practice, which OSCEs allow students to enter, is different in almost every respect. The subject is a practitioner whose object is to care for patients, not in some imagined future, but now. The tools are clinical procedures, instruments, drugs, and written guidelines, which are likely to differ not just between different hospitals and community settings but between individual wards and consulting rooms. There are official rules; but it may be unofficial rules, which contradict official rules, that enable new doctors to practise in fraught working environments (McLellan et al. 2015). The community now includes real patients, whose illnesses, responses to treatments, preferences, individual quirks, and capabilities to co-participate are absolutely non-standard. The division of labour involves working with doctors at all grades of seniority and from many different medical specialties, nurses, pharmacists, and other professionals, who may be almost as non-standard as patients. Working relationships are fluid and the negotiation of hierarchies follows unwritten rules that are also fluid and non-standard. The next section uses empirical data to illustrate some of these tensions and contradictions.

Box 1. Key methodological details of studies from which we draw empirical examples.

With ethical approvals from Queen's University Belfast (ref. 15.39) and National University of Ireland Galway College of Medicine, Nursing, and Health Sciences, we recruited a group of 35 OSCE stakeholders from ten institutions (undergraduate and postgraduate) across three countries to participate in a full-day workshop. These participants had a range of roles including direct participation (student candidates, examiners, and SPs) and 'behind the scenes' responsibilities (invigilators, question writers, administrators, and statisticians). The explicit aim of the workshop was to question and challenge OSCEs. The dataset came from group discussions 'triggered' by carefully constructed activities focusing on the present, then the past, then the future of OSCEs in HPE. AT served as an interpretive heuristic. We audio-recorded group discussions and transcribed them verbatim. Since some participants (particularly SPs) did not say much during these group activities, we conducted a second study to which we recruited seven SPs (none of whom had participated in the first study) for individual interviews. Participants in both studies provided informed and written consent. A dialogic approach to qualitative analysis informed by Sullivan (2012) guided our analysis of both group discussions and interviews, which went beyond the words participants uttered. We attended to extra-linguistic factors of the societal, cultural, and historical context, by examining how participants used social language. We paid specific attention to the appearance of tensions and contradictions in participants' speech acts.

Source of empirical data

Box 1 outlines key methodological features of two studies, from which we draw illustrative empirical data, comprehensive detail of which is available in thesis form (Reid 2018). We analysed participants' language using a critical discourse approach. Discourse contains traces of the cultural and historical origins of the present. Analysis of discourse can, justifiably, be used to inform the application of AT.

Standardised behaviour

As one student participant said: 'yes there is the aspect of listening to patients ... making sure you've explained well but because it's five minutes, in the back of your mind you're like, "try and get this patient out as quickly as possible." You know like you're trying to tick the boxes. And, you know, that really shouldn't be the way it is'. Clinical communication, in this example, is reified by tick-boxes on an examiner's checklist. The student acknowledges that using time efficiently to 'get the marks', rather than listening and explaining well, 'shouldn't be the way it is'. But OSCEs create a contradiction. The student will not qualify and will, therefore, be of no use to patients at all without getting marks that provide evidence of competence. Having to perform the most fundamental of all clinical skills in a formulaic way is a price that has to be paid for a medical qualification.

An SP participant expanded this contradiction: 'but if they're killing you nicely, they'll get a lovely mark!' A 'lovely mark' trivialises medical practice and reduces the OSCE station to a task for children rather than young professionals. An examiner expressed this similarly: 'Now the problem with tick boxes is that some students do a random spatter of questions and they tick these things at different points through the thing and they end up getting thirty-seven or thirty-eight ticked, but you hate the way they've done it. You give them a very poor score at the bottom. But they've got thirty-seven, thirty-nine, so even when the borderline regression is done, that student will still pass'. Despite examiners being experts (as senior clinicians) in clinical communication, the process of standardisation, as reified by the checklist, has more agency than the practitioner in determining whether a student should be allowed to practise. The clinician's ability to resist the power of standardisation was limited to referring dismissively to 'ticking these things at different points through the thing'. Again, formulaic behaviour to demonstrate competence trumps professional expertise.

In addition to the written rules of standardisation, unwritten rules determined students' behaviour. SPs noted how candidates did not actually listen to what was being said: 'you say your father died when he was in his 50s of heart disease and people going, "good, good, good, good" I'm seeing this happen!' This is a parody because actually listening to what patients are saying (as opposed to ticking a box marked 'hearing') is core to authentic clinical care but OSCE checklists were insensitive to the difference. Saying 'good good' conformed to the OSCE rule of demonstrating empathy, in exchange for which a checklist mark would be awarded, when the student had broken the most fundamental rule of empathic behaviour, which is to listen sensitively.

Another unwritten rule was that ‘patients’ in OSCEs were usually simulating disease without having the disease. Candidates, who were conditioned to expect normality, ‘looked’ (for which they earned a mark) without actually seeing. They might, with impunity, reel off ‘there are no scars’ even when the SP they were examining had a very obvious scar, which was unrelated to the ritualistic physical examination they had to demonstrate to earn marks.

An examiner further captured this tension with, ‘there’s the OSCE game, ok. So we just need to be aware that there is a sort of game going on. And that there is a sort of way of doing things’. The ‘ok’ in this utterance represented the speaker ‘laying his cards on the table’. A ‘game’ has rules, and the rules of the OSCE game are to behave in ways that are rewarded with marks in psychometrically valid calculations. These numbers meet the needs of regulatory accountability admirably well. They can be exchanged for a qualification but, as the examples illustrate, these numbers may have no use value and may even be harmful.

Tensions and contradictions

The OSCE paradigm of standardisation for reliable testing, which exchanges test scores for the right to practise, has unintended consequences over and above the adverse effects on clinical communication and physical examination described above. It is time with real patients that teaches students the shades of grey that make up illness, suffering, and wellbeing (see Bleakley 2020 in this issue). It is experiencing a wide variety of clinical presentations that teaches students about disease. And it is co-participating in practice that makes students capable clinicians. Students learn to care for patients by following the largely unwritten rules of behaving appropriately in clinical environments, using the tools of practice, and collaborating with peers, more senior doctors, and non-medical health professionals. All of this, which takes place in the activity system of authentic practice, is time-consuming and sometimes unrewarding for students.

Preparing for OSCEs is a very different activity. The mediating artefacts are performance checklists and the physical surroundings of libraries and coffee shops. Students use time efficiently in these settings to rehearse routines that maximise success on checklist scoring matrices rather than (from the OSCE standpoint) inefficiently and ineffectively in authentic practice. For medical students, the exchange value of being allowed to become a doctor exceeds the use value of being able to practise as a doctor.

The future: Harnessing tensions as possibilities for change

In envisaging a different future for OSCEs, we caution against losing sight of the positive benefits they have brought. OSCEs are a practical manifestation of a paradigm shift towards standardisation and reliability. Making that shift was relatively easy for decontextualised knowledge. Doing so for practical testing was more problematic. Our criticism is not of OSCEs *per se*; rather, we criticise using a tool that is incommensurate with practice to prepare students to practise. OSCEs have helped the field of medical

education to progress by introducing a (previously non-existent) practical component to assessment in some parts of the world. They have promoted practical skills, rather than just knowledge, in the earlier stages of curricula. When the goal is to prepare students for practice-based learning, rather than certifying them fit to practise, it makes sense to train and test competence in stable and standardised conditions using OSCEs.

The paradigm of standardisation and assessment has, though, shown signs of fracturing. Hodges posed searching questions about the performativity and reliance on psychometrics and production that characterise OSCEs (Hodges 2009). Our AT analysis complements Hodges’ Foucauldian interpretation by critiquing the historicity and rise to dominance of OSCEs. We have shown multivoiced and contradictory aspects of OSCEs in the present. We have identified contradictions and tensions, which could be harnessed to expand the activity of medical education. Transformations of activity are not linear processes with fixed destinations. The historical development of activities opens up possible spaces as zones of proximal development (ZPDs). These are contested spaces converging around the development and expansion of an object, which, in the case of HPE, is patient care. The discussion that follows gazes into the future of assessment practice by considering three ‘spearheads’ of expansion: lessening of tensions, expansive learning, and knotworking.

Lessening tensions

Stakeholder participants in our research identified tensions around patient care being the object of assessment activity. Patients are often absent from OSCEs and represented by actors adhering to standard scripts. One (student) participant highlighted that ‘the patients in front of us will all adhere to the formula that we have ... so we don’t actually have to use any skills in changing our approach and things because everyone’s the same!’ If ecological validity rather than just psychometric reliability were the dominant rule of the assessment activity system, involving real patients could reorient the activity towards the object of caring for patients. The next section describes programmatic assessment, which can relatively easily involve real patients, though OSCEs can do so too. Indeed, the UK GMC has made real patient involvement a requirement for the (OSCE based) clinical component of the soon to be implemented NLE. It will be important, though, to ensure that authentic patient involvement does not itself come under pressure to standardise.

Expansive learning

Engeström describes expansive learning as a lengthy journey across the ZPD. Adopting programmatic assessment rather than ‘single point in time’ decontextualised OSCEs could be a step in such a cycle of expansion. The programmatic approach uses multiple low stakes assessments in different contexts to provide learner profiles; these contexts could easily be workplaces, where the activity system of authentic patient care is dominant (Van der Vleuten et al. 2012). Participants in our studies looked to ways of embedding future assessments in practice settings, where rules of

uncertainty and complexity prevail, rather than the rules of certainty and simplicity that dominate OSCEs. It will be hard to change assessments that are stuck in a paradigm of standardisation and stability. Yet this can be done. Published assertions that 'reductionism is not the only way to ensure rigour in high stakes assessment' and, 'standardisation is not the only route to equity' (Schuwirth and Ash 2013) highlight a reorientation towards programmatic assessment approaches.

Knotworking

Caring for patients is not an individual task. Healthcare is being conceptualised, increasingly, as a team activity rather than an individual pursuit (Lingard 2012). Today's healthcare tends to involve complex collaborations between fluidly constructed and frequently changing teams, working over time and place. The companion article by Varpio and Teunissen (2020) characterises interprofessional healthcare teams as a quintessential example of another AT-derived concept: knotworking. Whilst this is how we work in healthcare; it is not how we assess learners. OSCEs are a largely individual exercise. Attempts to bring 'team' elements to OSCEs have largely failed to gain traction (Marshall et al. 2008). Growing acceptance of knotworking as the reality of work in the health professions has potential to drive innovation in assessment, perhaps through further cycles of expansive learning.

Conclusions

AT-informed research has highlighted tensions around OSCEs, explored assessment's ZPD, and suggested expansive possibilities for the future. There are some moves towards assessments, which are continuous, frequent, and ideally conducted within the activities of workplaces. A recent reflection by Holmboe characterised moves in this direction as 'a paradigm shift struggling to be realised' (Holmboe 2018). Struggle around paradigmatic transitions is an inevitability. Kuhn himself noted that 'when an individual or group first produces a synthesis able to attract most of the next generation's practitioners, the older schools gradually disappear. In part their disappearance is caused by their members' conversion to the new paradigm' (Kuhn 1962). OSCEs might soon come to be such an 'older school' as practitioners transition – convert – to assessments in naturalistic, authentic settings where caring for real patients is the core activity.

Acknowledgements

Thanks to Dr. Mairead Corrigan, Professor Pascal McKeown, and Professor Peter Cantillon for supporting the research on which this article draws.

Disclosure statement

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

Notes on contributors

Helen Reid, BM, PhD, is a GP. Her research interests concern assessment and healthcare workplace learning (particularly in community contexts), drawing on a range of critical methodologies.

Gerard J. Gormley, MB, MD, is a GP and Professor in Simulation. He has a particular interest in education scholarship – especially relating to simulation.

Tim Dornan, BM, PhD, is Professor of Medical Education at QUB and Emeritus Professor, Maastricht University, The Netherlands. He is an internist and endocrinologist, whose main interest is workplace learning. His methodological interests include dialogic discourse analysis and a range of qualitative methodologies inspired by sociocultural theory and phenomenology.

Jennifer L. Johnston, BM, PhD, is a critical educationalist and practising GP. Her interests lie in uncovering implicit power dynamics, addressing inequalities through education, and raising critical consciousness.

ORCID

Helen Reid  <http://orcid.org/0000-0001-8530-1766>

Gerard J. Gormley  <http://orcid.org/0000-0002-1701-7920>

Tim Dornan  <http://orcid.org/0000-0001-7830-0183>

Jennifer L. Johnston  <http://orcid.org/0000-0002-3999-8774>

References

- Bleakley A. 2020. Embracing ambiguity: curriculum design and activity theory. *Med Teach*. DOI:10.1080/0142159X.2020.1795099.
- Engeström Y, Pyörälä E. 2020. Using activity theory to transform medical work and learning. *Med Teach*. DOI:10.1080/0142159X.2020.1795105.
- Eva K, Rosenfeld J, Reiter H, Norman G. 2004. An admissions OSCE: the multiple mini-interview. *Med Educ*. 38(3):314–326.
- Freidson E. 1970. *Profession of medicine: a study of the sociology of applied knowledge*. New York (NY); London: Harper and Row.
- Goodwin D. 2018. Cultures of caring: healthcare 'scandals', inquiries, and the remaking of accountabilities. *Soc Stud Sci*. 48(1):101–124.
- Harden RM, Gleeson FA. 1979. Assessment of clinical competence using an Objective Structured Clinical Examination (OSCE). *Med Educ*. 13(1):41–54.
- Harden RM, Stevenson M, Downie WW, Wilson GM. 1975. Assessment of clinical competence using objective structured examination. *Br Med J*. 1(5955):447–451.
- Hodges B. 2009. *The objective structured clinical examination; a socio-history*. Köln, Germany: Lambert Academic Publishing AG & Co.
- Holmboe ES. 2018. Competency-based medical education and the Ghost of Kuhn: reflections on the messy and meaningful work of transformation. *Acad Med*. 93(3):350–353.
- Kuhn TS. 1962. *The structure of scientific revolutions*. 2nd ed. Chicago (IL): The University of Chicago Press.
- Lingard L. 2012. Rethinking competence in the context of teamwork. In: Hodges B, Lingard L, editors. *The question of competence*. New York (NY): Cornell University Press; p. 42–69.
- Lesky E. 1970. The development of bedside teaching at the Vienna medical school. In: O'Malley CD, editor. *The history of medical education: an international symposium held February 5–9, 1968*. Berkeley; London: University of California Press.
- Marshall D, Hall P, Taniguchi A. 2008. Team OSCEs: evaluation methodology or educational encounter? *Med Educ*. 42:1111–1146.
- McLellan L, Yardley S, Norris B, De Bruin A, Tully M, Dornan T. 2015. Preparing to prescribe: how do clerkship students learn in the midst of complexity? *Adv Health Sci Educ Theory Pract*. 20(5):1339–1354.
- Ponnamperuma GG, Karunathilake IM, McAleer S, Davis MH. 2009. The long case and its modifications: a literature review. *Med Educ*. 43(10):936–941.
- Reid H. 2018. *Examining OSCEs: critically exploring discordant voices*. Belfast: Queen's University Belfast.
- Rotgans JI. 2012. The themes, institutions, and people of medical education research 1988–2010: content analysis of abstracts from six journals. *Adv Health Sci Educ Theory Pract*. 17(4):515–527.

- Schuwirth L, Ash J. 2013. Assessing tomorrow's learners: in competency-based education only a radically different holistic method of Assessment will work. Six things we could forget. *Med Teach*. 35(7): 555–559.
- Sullivan P. 2012. *Qualitative data analysis using a dialogical approach*. London: Sage.
- Van der Vleuten CPM, Schuwirth L, Driessen E, Dijkstra J, Tigelaar D, Baartman L, Van Tartwijk J. 2012. A model for programmatic assessment fit for purpose. *Med Teach*. 34(3):205–214.
- Varpio L, Teunissen P. 2020. Leadership in interprofessional healthcare teams: empowering knotworking with followership. *Med Teach*. DOI:10.1080/0142159X.2020.1791318.