An Accidental Career

Sydney M. Smee

Abbreviations

ACIR   Arizona Clinical Interview Rating scale
AMEE   Association of Medical Educators of Europe's
ASPE   Association of Standardized Patient Educators
ATLS   Advanced Trauma Life Support
IMSH   International Meeting on Simulation in Healthcare
OSCE   Objective Structured Clinical Examination
PI   Patient Instructors
SOBP   Standards of Best Practice
SP   Standardized/Simulated Patient
SPE   Standardized Patient Educator

It was 1973, I was 16 years old, and I was waiting to see a doctor. I kept going over and over what I would say, how to explain being pregnant and scared. When the doctor came in, I realized he was as anxious as I was; probably because his colleagues were watching us through the two-way mirror. I was simulating a role, but he was not. He would be receiving feedback about his performance. Suddenly I was not so nervous. I was doing the simulation as a replacement for my sister who had signed up to do it and then could not make it. The whole experience was a fascinating beginning to an accidental career.

Dr. Howard Barrows was introducing simulated patients into the health sciences curriculum at McMaster University, a new medical school close to where I lived. Gayle Gliva-McConvey was the SP trainer and she was the one who taught me the most about being a simulated patient or SP. Later she coached me in training others to be SPs. Most of my early work at McMaster involved simulating for small group teaching sessions. Over 12 years, I learned to simulate many patient problems and in doing so, I also learned a bit of medicine, acquired some medical terminology and found out quite a bit about history taking and physical examination techniques.

SPs create powerful learning moments. One time I was presenting with a total lack of lower limb sensation or movement as part of a presentation of multiple sclerosis. The occasion was a small group teaching workshop and I was assessed by a faculty volunteer. As he examined me, his sensation testing became rather aggressive. He kept pushing a pin deeper into my legs and feet, trying to elicit a response. Afterwards the facilitator led a group discussion providing him with feedback and discussing small group teaching techniques. When the session was over, I stood up. The volunteer went pale. He had come to believe that I was a real patient and that I had not felt anything because he could not elicit a pain response. My discomfort was worth it. He had forgotten it was a simulation and fully engaged in the learning process. On another occasion I was lying limp on a stretcher, supposedly only semi-conscious, during an Advanced Trauma Life Support (ATLS) course. The physician was preparing to log roll me away from him; which was an unsafe maneuver and would likely cause me to fall off the stretcher. I knew that if I stayed limp and fell, he would never make this mistake again. I wondered if I should do it. We were never expected to risk injury as an SP but our goal as SPs was to make each simulation as authentic as possible. I think I was willing to roll off that stretcher to maintain the simulation. Fortunately, I didn't have to. Part way into the maneuver the physician realized his mistake. I believe that figuring it out himself was an important learning moment and I was glad I had stayed in role long enough for it to happen.

My part-time job as an SP saw me through high school, supported me while I completed an undergraduate degree in political science and supplemented my income as I worked at other jobs. Then, for a short while I covered for Gayle during her maternity leave. During that time, Dr. Paula Stillman called the program, hoping to recruit Gayle Gliva-McConvey to join her at the University of Massachusetts. Gayle said no

S. M. Smee (✉)
Medical Council of Canada (ret), Ottawa, ON, Canada

but suggested I apply. I did and shortly found myself living in Worcester and working as the coordinator for Dr. Stillman's Patient Instructor program.

I had never heard of patient instructors, although I had been a gynecological teaching associate for several years at McMaster. I quickly learned that medical students would meet one-on-one with a series of patient instructors to take a history or to complete a physical examination. The patient instructors used their own medical history and findings, and afterwards provided feedback to the student about their basic clinical skills. Patient instructors were required to complete a training program that introduced them to the basic physical exam techniques, basic history taking skills, and to score some very detailed checklists along with the Arizona Clinical Interview Rating scale (ACIR). What I had learned about clinical skills at McMaster had been by osmosis over 12 years of simulation. I quickly realized that I needed more formal knowledge of physical exam techniques and history taking skills. Thankfully, I was granted permission to take the practical component of the Year Two clinical skills course with the medical students. I was more self-taught when it came to coaching the Patient Instructors (PIs) with video-based exercises to promote reliable scoring. However, my years of being an SP for small group teaching sessions and my training work from a volunteer organization informed how I facilitated these training sessions.

I liked working with the PIs, but I found the detailed checklists rigid and constraining. This was a very different approach to what I knew from patient simulation and providing feedback on interactions from a patient-based perspective. Patient Instructors commonly used their own histories and provided feedback on specific skills. They did not need to learn a role, but they did benefit from learning how to present their cases without leading the medical students and learning to present their story as fresh, even after many repetitions.

As part of my work, I assisted with a large-scale research study that examined the value of using standardized patients to assess the clinical skills of residents across multiple New England training programs. The term patient instructor was replaced by the new term because the focus was on assessment of skills, not on providing feedback. Now SP meant something a bit different. My contribution earned me third authorship on the paper that reported on this study [1]. While I appreciated the acknowledgement, I did not understand its career value until much later. I didn't know I was on a career path.

After 2 years, I returned to Canada. I knew assessment work was important but did not see it as being my long-term focus. Professionally speaking, I went on a "walkabout". I did small contracts, I travelled, and then I became the coordinator for a hospice volunteer program. My experience with the patient instructor program was highly transferable. I believed I was on a career path. However, to stay on that path and maybe become a program director at a larger institution, I needed more education. Back to school I went. I registered in a Master of Education program with a special interest in Adult Education.

While pursuing my degree and looking for a new position, I received a phone call. Would I be interested in a 3-year project to develop a high-stakes clinical skills assessment for the Medical Council of Canada? They were looking for a standardized patient (SP) trainer. I had never heard of the Medical Council of Canada and somehow forgot that assessment did not interest me that much. Next thing I knew, I was part of a small team tasked with developing and piloting a 20-station Objective Structured Clinical Examination (OSCE). Not only was I unfamiliar with the Medical Council of Canada, I was also uninformed about OSCEs.

I quickly learned that an OSCE relies on the standardized presentation of a series of patient problems to ensure that a cohort of trainees is assessed against the same set of cases or test items. The fairness and objectivity of an OSCE is further enhanced by pre-set scoring criteria, most often in the form of detailed checklists. OSCEs rely on standardized patients (SPs) to present patient problems realistically and they require SPs to align their presentation with detailed checklists to ensure score reliability. I learned over time that these two objectives do not always coexist comfortably. By the time I was introduced to the OSCE at the Medical Council of Canada, there was a growing body of evidence to support piloting an OSCE for national licensure [2–12]. The pilot had three sites, each running multiple tracks of 20 stations [13]. Multiple SPs were presenting the same role at each site and across sites. Sixty patient cases were needed for the pilot and the anticipated first administration.

When I started the OSCE design had been determined but the content, the patient cases, had yet to be developed. Scoring would be done by physicians who would observe and score the examinees within each station. We were building something new from the ground up. We were creating training materials for SPs, for site staff, and for the examiners. There were formatting and production issues to solve; scoring processes to create, and budgets to manage; the task list was endless, the learning curve was steep.

My roots were in patient simulation. Being the SP trainer and later the manager for a national high stakes OSCE meant a growing distance from direct SP-related work. With time, the two reports that became most important to me were the annual budget and the post-exam analysis. Dollars and data were my measures of success. A three-year contract had become a long-term position. My director and mentor, Dr. David Blackmore, pushed me to go back to school. The Medical Council of Canada would allow me to continue working and somehow, despite saying no, I ended up in a doctoral program in education with a focus on measurement and test theory.

During 8 years of working and studying I thought a lot about how an OSCE is scored and how that might be improved. Perhaps the biggest criticism of OSCEs (other than their cost) is that short stations and detailed checklists deconstruct what it means to be a clinician [14–17]. A physician does not ever just examine a knee, they examine a patient with a knee problem. OSCEs that rely on checklists arguably promote the wrong kind of learning. Many medical trainees engage in rote performance. At each OSCE station they ask and do as many things as they can from generic, memorized lists to gain as many marks as possible, as easily as possible. Candidates provided me with examples of this kind of rote performance when they spoke with me about their results. I was assured by one candidate that he had been empathetic during the OSCE; he had taken a course and he knew that empathy equaled touching the patient's arm three times. Other candidates argued that they "had done everything". Why had they done poorly? They meant they had done everything on their generic checklist. These are test-taking behaviors, not a true demonstration of clinical skills and an unintended negative consequence of scoring OSCEs with checklists.

Short stations and detailed checklists also deconstruct patient simulation, beginning with SP training. For example, SP trainers need to know how to standardize SP responses to open-ended questions. There are at least 3 different strategies to help SPs provide naturalistic responses to open-ended questions without giving away too much information and thereby forcing the medical trainee to use follow-up questions. One is providing only one new piece of information, a second is repeating information already provided, including simply repeating the chief complaint, and a third is providing extraneous information to the question. However, the strategy that trainers' default to is training SPs to respond to an open-ended question with a question. So, when the SP is asked "What can you tell me about your foot pain? the SP responds with "What do you mean" or "Like what?" Candidates are forced to ask, "Is it sharp or dull?" "Does it throb?" "When did it start?" SPs answering a question with a question also promotes test-taking behaviors rather than rewarding good clinical performance.

Some trainers focus on unnecessary details in the pursuit of standardization. Once I was asked for the names of the patient's siblings. The siblings were peripheral to the patient's problem; standardizing the names did not matter. The trainer was striving to do a good job but was wasting time on details that were not critical to generating reliable scores.

On another occasion, I observed SPs being trained to present delirium. The SPs were to look around the room *about* four times during a 5-minute history. These SPs did look around at exactly 4 points during the practice, each time between questions from the physician. They gave a very mechanical presentation of a delirious, distracted patient. Then there are SPs who are accurate but sound scripted. "How would you rate your pain on a scale of 1 to 10 where 10 is the worst pain you can imagine?" "Seven." Instant reply. Not the more natural response of pausing slightly and then replying, "I don't know, it's bad, it's probably a seven."

These are examples of the erosion in authenticity that comes from standardizing SPs to a checklist. They are also examples of the impact, often negative, that OSCEs have had on SP trainers and SP educators. Standardization does matter and generating reliable scores when multiple SPs are presenting the same case requires clear case protocols. A key component of strong OSCE case writing is including fixed guidelines for SPs: "Only ask this question after 4 minutes" or "groan 3-4 times over 5 minutes" and "One answer only for each checklist item". The key to fair testing is that everyone sees the same cases so all the SPs doing the same case should be the "same", or at least as much the same as possible. However, SPs also need to align their responses to the questions and attitudes of each medical trainee, while still following the protocol for their case. When this nuance is lost, the best of what SPs bring to clinical assessment is undermined. When training approaches and the use of SPs are defined narrowly, as they are through an OSCE lens, then the full scope of SP-based educational activities is underdeveloped. SPs are wonderful teachers and powerful adjuncts to clinical faculty. They can provide direct, constructive feedback to learners about communication, history taking and basic physical exam skills in a variety of contexts. The introduction of simulated patients made OSCEs possible and OSCES have advanced the use of standardized patients in medical education. However, there is a tension that exists between patient simulation and high stakes assessment, between authenticity and reliability, that leading SP educators are always managing.

No More Accidents

No one grows up dreaming of becoming an SP educator. More often, individuals come to the field from a variety of backgrounds. They bring with them different areas of expertise that need to be adapted, expanded and integrated into a new field of practice. The requisite knowledge base encompasses everything from best practices in simulation to a grounding in educational and assessment principles. Expected skills range from teaching and coaching to human resource and program management skills. The Association of Standardized Patient Educators (ASPE) Standards of Best Practice (SOBP) [18] define the scope of required knowledge and skills and are an essential resource to aspiring SP educators. I remember the need for standards being raised by Gayle Gliva-McConvey at the 1993 *Set the Standard*

conference for SP educators (SPEs) in Calgary, Alberta, Canada. Twenty years later I was included in a working group of SP educators she convened in Vero Beach Florida. Gayle insisted that we could and would draft practice standards for SP educators. We did. ASPE leaders saw that work through to publication. The practice standards challenge all SP educators to look at their own practice and their own programs with clear eyes, to reflect on where to focus their professional development, and to advocate for SPs within their own institutions. The practice standards are a framework that represent the best of 5 decades of development in our field and are a guide to the SP educator community of practice as they meet the future.

The Standards of Best Practice [18] define the scope of SP educator practice, but they do not define a career path. The challenge for each individual is to create their own apprenticeship; an apprenticeship tailored to their individual context, an apprenticeship that respects their unique expertise and that addresses where they need to grow. Understanding the limitations of self-assessment [19] and learning about self-directed assessment [20] may be particularly empowering for SP educators who are creating their own path of professional development. Self-directed assessment seeking is a self-driven process of looking outward, not inward, and seeking feedback to guide and promote performance improvements. The informed self-assessment model proposed by Sargeant and her colleagues captures a complex process in five interactive components: (1) sources of information, (2) interpretation of information, (3) responses to information, (4) external and internal conditions that influence the first three steps, and (5) the tensions created by competing internal and external factors. First is information, that can come from external processes such as a course, or it can come from people, such as one's peers, co-workers, and supervisors. Information can also come from one's emotional and internal states. Next, information is interpreted through reflection, calibrating it against other feedback, and filtering it. We may accept or ignore information that does not fit with what we believe, or we may reject and then consider it, leading to further reflection and even acceptance of it. Information that confirms how we see ourselves is often simply accepted, only sometimes questioned. How we interpret and respond to information is influenced by the context in which we receive the information, our relationships with others, how we judge the credibility of the source, and our personal attributes, like our emotions and our curiosity. This whole process creates and is moderated by tensions; such as the wish to perform better versus the wish to appear informed and competent to others or the wish of the other person to give us genuine feedback versus their wish to simply validate positive attributes and avoid more uncomfortable conversations. Their tension is mirrored by our own wish for genuine feedback versus our fear of disconfirming and discomforting information.

Understanding the need for meaningful input from others and the conditions needed to elicit it, is an invaluable underpinning to having an intentional career. The scope of knowledge and skill required of even a new SP educator today means that accidental careers are less possible than it was during the early years. However, the resources available to SP educators are far greater. ASPE is an expanding community of practice that comes together at the annual ASPE meeting to share expertise and to promote good practice. ASPE has many experts within its membership who have developed critical resources; including the literature reviews and the research database of all things SP developed and made available by Karen Szauter; there is the textbook, *Simulated Patient Methodology: Theory, Evidence and Practice*, edited by Debra Nestel and Margaret Bearman [21]; there is Peggy Wallace's book *Coaching Standardized Patients for Use in the Assessment of Clinical Competence* [22], and there is *Objective Structured Clinical Examinations: 10 Steps to Planning and Implementing OSCEs and Other Standardized Patient Exercises* [23], edited by Sondra Zabar, Elizabeth Kachur, Adina Kalet and Kathleen Hanley. The International Meeting on Simulation in Healthcare (IMSH), the biennial Ottawa Conference on assessment of clinical competence, and the Association of Medical Educators of Europe's (AMEE) annual medical education conference all have much to offer SP educators, just as SP educators have much to offer at these meetings.

Looking Ahead

Discovering research in cognitive psychology that focused on clinical assessment jolted me out of a certain complacency about OSCE design and OSCE scoring [24–28] I was challenged to think about the cognitive load of the rating task, the impact of first impressions on raters, the narrative nature of social judgments, and how to align the language on scoring instruments with how raters think. Their research raises questions. "Can we shorten checklists and still have reliable scores?" "Will making the cognitive load less minimize biases like first impressions?" "Can we design checklists and rating scales that reflect how raters think rather than trying to train raters to think like test developers?" "Should there be two raters – scoring different aspects of the same performance?" If there were, there would be more data and that usually means more reliable scores.

In my own practice, a new blueprint at the Medical Council of Canada [29] challenged the test committee and the OSCE team to develop more authentic, complex cases that would assess more than the basic clinical skills of post-

graduate trainees. Success would require scoring strategies that did *not* reward the rote performance so often seen in OSCEs. Detailed checklists would not work in this context.

Checklists are useful tools, but they are best suited to scoring when thoroughness matters and for assessing beginner levels of ability or procedural tasks. They are useful when the time for rater training is limited or the time available for the marking task is limited. Rating scales are often promoted as an antidote to checklists. Rating scales are best suited to scoring behaviors, aspects of performance that are "more or less" done, and for capturing increasing levels of expertise or judgment. However more time is needed for rater training and for the rating task than is true for checklists. An early and often cited study [30] showed that rating scale scores were more reliable and discriminated better across levels of expertise, but the authors cautioned that the rating scales might have been too generic. Further, the raters in the study scored both checklists and rating scales which confounded the reliability analysis of the rating scale data. Did the checklists help standardize the raters before they completed the rating scale? Also, the checklists were designed to assess medical students, but the study compared the performance of different levels of post graduate trainees and experienced physicians. Was the issue the checklist format or the student-focused content of the checklists? More recently the checklist versus rating scale debate has given way to using some combination of checklist and rating scale items, an approach that is increasingly seen as best practice [31].

Without the constraint or framework of detailed checklists, the SPs and SP educators will need to use far more judgment to ensure that the kind of cases that Medical Council of Canada is developing are presented reliably. The SP training shortcuts of the past few years will be insufficient to support this kind of new content. SP trainers who are stuck in a paint-by-numbers approach will need to develop new insight and skills. However, these SP trainers are stuck because of high workloads, only knowing how to train for OSCE cases, or because they have not had enough training and support to know what is possible. I believe that achieving greater authenticity within an OSCE framework is possible if SP trainers have the necessary support and if they have the strong SP training skills and the good judgment that comes from an understanding of the underlying assessment principles. There are already many SP educators, working within their institutions, who are collaborating on SP-based innovations and promoting excellence in learning and assessment. I also believe the drive for more authentic and complex cases and the concomitant challenge to SP educators is not unique to the work at the Medical Council of Canada.

There are limits to what can reasonably be simulated in an OSCE, especially in terms of physical signs and symptoms. Even in educational exercises there are limitations. Simulated patients are not actual patients. That is a constraint and a strength. Trying to figure out more and fancier ways to create simulations in the OSCE or ways to overcome the physical limitations of SPs does not seem like the best strategy to me. Finding better ways to train and coach SPs on what they can do best seems far more important. However, some of what SPs do best are also the things that are hardest to standardize. Emotional roles are one example; more interactive roles are another. Basic history-taking and physical exam roles are driven by the trainees, so these roles are primarily reactive and are more easily simulated, more easily scored.

Interactions driven by the SP require more judgment from the SP, there is room for more variance. Interactive roles include patients questioning how their problem is being managed, patients who present ethical challenges, and SPs who simulate clinical colleagues demanding some form of response from the trainee are a small sample of a wide range of complex roles that will require a new understanding of 'standardized'. Some of these more complex presentations are being well explored within SP programs. Learning from these educational initiatives should and can inform what is possible in assessment, even within the restrictions of high stakes OSCEs.

Final Reflection

An accidental career was more possible 30 and 40 years ago. SPs in medical education were an innovation, OSCEs were new; everyone was learning. In many ways my accidental career evolved as the field itself evolved. I was fortunate to work with leaders in the field and to be a part of the Medical Council of Canada for over 25 years. I benefited tremendously from rich, if unintended, learning opportunities. First were my years as a simulated patient in a problem-based curriculum, where I learned some medicine and I learned about teaching. Later, years of working with test committees and clinical case writers taught me even more about medicine and a lot about assessment. I was blessed with mentors who fostered my learning and who gave me increasingly responsible roles that allowed me to grow, to experiment, to lead. Intentional learning, my post-graduate education, deepened my understanding of critical knowledge and broadened my perspective but came late in the process.

Today, there is a maturing community of practice, a large body of research and reference materials. I do not believe that an accidental career is as possible. One may still enter the field "accidentally" since many SP educators still come from other fields. However, I think there is an onus on today's SP educator to be intentional in their professional development; to understand and assimilate what has already been learned and accomplished so they can build from it, not recreate it.

References

1. Stillman PL, Swanson DB, Smee SM, Stillman AE, Ebert TH. Assessing the clinical skills of residents with standardized patients. Ann Intern Med. 1986;105(5):762–71.
2. Vu NV, Steward DE, Marcy M. An assessment of the consistency and accuracy of standardized patients' simulations. J Med Educ. 1987;62:1000–2.
3. Newble DI, Swanson DB. Psychometric characteristics of the objective structured clinical examination. Med Educ. 1988;22:325–34.
4. Colliver JA, Verhulst SJ, Williams RG, Norcini JJ. Reliability of performance on standardized patient cases: a comparison of consistency measures based on generalizability theory. Teach Learn Med. 1989;1(1):31–7.
5. Swanson DB, Norcini JJ. Factors influencing reproducibility of tests using standardized patients. Teach Learn Med. 1989;1:158–66.
6. van der Vleuten CPM, Swanson DB. Assessment of clinical skills with standardized patients: state of the art. Teach Learn Med. 1990;2:58–76.
7. van Luijk SJ, van der Vleuten CPM, editors. A comparison of checklists and rating scales in performance-based testing. In: Current developments in assessing clinical competence; 1990. Montreal, Can-Heal Publications Inc.; 1992.
8. Vu NV, Colliver JA, Verhulst SJ, editors. Factor structure of clinical competence as assessed in a performance-based examination using standardized patients. In: Current developments in assessing clinical competence; 1990. Ottawa/Montreal: Can-Heal Publications; 1992.
9. Cohen R, Rothman AI, Poldre P, Ross J. Validity and generalizability of global ratings in an objective structured clinical examination. Acad Med. 1991;66:545–8.
10. Tamblyn R, Klass D, Schnabl G, Kopelow M. The accuracy of standardized-patient presentations. Med Educ. 1991;3:74–85.
11. Grand'Maison P, Lescop J, Rainsberry P, Brailovsky CA. Large-scale use of an objective structured clinical examination for licensing family physicians. Can Med Assoc J. 1992;146:1735–40.
12. Parker-Taillon D, Cornwall J, Cohen R, Rothman AI, editors. The development of a physiotherapy national examination OSCE. The 5th Ottawa Conference: Approaches to the assessment of clinical competence; 1992. Ottawa, Canada 1992.
13. Reznick RK, Smee SM, Rothman AI, Chalmers A, Swanson DB, Dufresne L, et al. An objective structured clinical examination for the licentiate: report of the pilot project of the Medical Council of Canada. Acad Med. 1992;48:487–94.
14. JPW C, Neville AJ, Norman GR, editors. The risks of thoroughness: reliability and validity of global ratings and checklists in an OSCE. Advances in medical education: proceedings of the seventh Ottawa international conference on medical education; 1997. Dordrecht: Kluwer Academic Publishers; 1997.
15. Hodges B, McNaughton N, Regehr G, Tiberius RG, Hanson M. The challenge of creating new OSCE measures to capture the characteristics of expertise. Med Educ. 2002;36:742–8.
16. Nichols PD, Sugrue B. The lack of fidelity between cognitively complex constructs and conventional test development practice. Educ Meas Issues Pract. 1999;18(2):18–29.
17. Norman GR. Editorial – checklists vs. ratings, the illusion of objectivity, the demise of skills and the debasement of evidence. Adv Health Sci Educ. 2005;10(1):1–3.
18. Lewis KL, Bohnert CA, Gammon WL, Hölzer H, Lyman L, Smith C, et al. The association of standardized patient educators (ASPE) standards of best practice (SOBP). Adv Simul. 2017;2(1):10.
19. Eva KW, Regehr G. "I'll never play professional football" and other fallacies of self-assessment. J Contin Educ Health Prof. 2008;28(1):14–9.
20. Sargeant J, Mann K, Van der Vleuten C, Metsemakers J. "Directed" self-assessment: practice and feedback within a social context. J Contin Educ Health Prof. 2008;28(1):47–54.
21. Nestel D, Bearman M, editors. Simulated patient methodology: theory, evidence and practice. Hoboken New Jersey: Wiley Blackwell; 2014. p. 168.
22. Wallace P. Coaching standardized patients: for use in the assessment of clinical competence. New York: Springer Publishing Co; 2007.
23. Zabar S, Kachur E, Kalet A, Hanley K, editors. Objective structured clinical examinations: 10 steps to planning and implementing OSCEs and other standardized patient exercises. New York: Springer; 2013.
24. Crossley J, Johnson G, Booth J, Wade W. Good questions, good answers: construct alignment improves the performance of workplace-based assessment scales. Med Educ. 2011;45(6):560–9.
25. Gingrich A, Regehr G, Eva KW. Rater-based assessments as social judgments: rethinking the etiology of rater errors. Acad Med. 2011;86(86):S1–7.
26. Tavares W, Eva KW. Exploring the impact of mental workload on rater-based assessments. Adv Health Sci Educ Theory Pract. 2013;18(2):291–303.
27. Wood TJ. Exploring the role of first impressions in rater-based assessments. Adv Health Sci Educ. 2014;19(3):409–27.
28. Yeates P, O'Neill P, Mann K, Eva KW. Seeing the same thing differently: mechanisms that contribute to assessor differences in directly-observed performance assessments. Adv Health Sci Educ Theory Pract. 2013;3:325–41.
29. Touchie C, Streefkerk C. Blueprint project – qualifying examinations blueprint and content specifications. Medical Council of Canada: Ottawa Canada; 2014.
30. Hodges B, Regehr G, McNaughton N, Tiberius RG, Hanson M. OSCE checklists do not capture increasing levels of expertise. Acad Med. 1999;74(10):1129–34.
31. Swanson DB, van der Vleuten CPM. Assessment of clinical skills with standardized patients: State of the art revisited. Teaching and Learning in Medicine. Int J. 2013;25(sup1):s17–25.