Pharmacy Education, 2017; 17 (1) 73 - 80

Qualitative evaluation of a cumulative exit-from-degree objective structured clinical examination (OSCE) in a Gulf context

AHMED HESHAM SOBH [1], MOHAMED IZHAM. M.I.[1], MOHAMMAD I. DIAB[1], SHANE A. PAWLUK[1], ZUBIN AUSTIN[2], KYLE J. WILBY[1*]

[1]*College of Pharmacy, Qatar University, Doha, Qatar*

[2]*Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario, Canada*

Abstract

Introduction: Cultural adaptation of assessments into new contexts requires extensive evaluation to ensure validity of examination methods.

Aims: The study objective was to critically analyse a cumulative OSCE for graduating pharmacy students in Qatar from stakeholder perspectives.

Methods: Two focus groups were conducted with stakeholders to perform a SWOC (strengths, weaknesses, opportunities, challenges) analysis. Discussions were recorded, transcribed, and analysed for content using an inductive analytical approach.

Results: The content analysis resulted in 20 categories. Strengths included training, assessment practices, familiarity, standardisation, and satisfaction. Weaknesses included discomfort, assessment practices, exam organisation, and training. Opportunities included future licensure, regulator support, improvement, and standardised actor pool. Challenges included novelty, failure policies, specialised pharmacists, preparation of practicing pharmacists, collaboration, cultural differences, and OSCE overall scoring.

Conclusion: Adaption of a cumulative OSCE in Qatar is positively perceived by stakeholders yet opportunities exist to improve design and implementation within our context.

Keywords: *OSCE, Evaluation, Pharmacy Education, Medical Education*

Introduction

Performance-based assessment, also known as a type of authentic assessment (Wiggins, 1989), requires students to demonstrate achievements in learning by performing tasks resembling real practice (Gronlund, 1998; McMillan *et al*., 2007). The purpose of this assessment is to simulate real life situations and integrate "higher-order thinking skills" necessary for success in real life situations (Swanson *et al*., 1995; Gronlund, 1998; McMillan *et al*., 2007). These attributes allow students to integrate knowledge and skills to demonstrate effective performance on simulated real-life tasks (Wiggins, 1993). Therefore, performance-based assessments are very important in health professional training, where students develop skills to be competent in various professional roles.

Objective Structured Clinical Examinations (OSCEs) are a type of performance-based assessment that requires learners to show their skills and abilities in a simulated environment. This type of assessment is currently considered a model performance-based assessment for health professions worldwide. Although OSCEs are not necessarily new to the Middle East (Zahid *et al*., 2011; Karim *et al*., 2012; Selim *et al*., 2012; Turan & Konan, 2012; Raheel & Naeem, 2013; Abdelaziz *et al*., 2015), few schools have adapted and used OSCEs in the field of pharmacy as a part of their curricula (Johnson *et al*., 2011; Al-Azzawi *et al*., 2013). At the Qatar University College of Pharmacy, the OSCE was adapted from the Canadian context in 2014. It was successfully constructed, implemented and organised as a final cumulative assessment for graduating pharmacy students. Upon the second iteration, the OSCE design and implementation processes must be comprehensively studied in order to improve its adaptability into this new context and to offer insight for other global institutions attempting to incorporate high stakes cumulative OSCEs into their own curricula.

This leads to the primary research questions of this study, which were, "What are stakeholder's perceptions of quality, utility, and value of a cumulative OSCE, and how can quality be improved to ensure uptake and

[*]Correspondence: Kyle John Wilby, *PO Box 2713, College of Pharmacy, Qatar University, Doha, Qatar. Tel: +974 4403 5606; Fax: +974 4403 5551. Email: kjw@qu.edu.qa*

ISSN 1447-2701 online

applicability for high stakes assessment?" In order to answer these questions, the specific objective of this study was to critically analyse the second iteration of a cumulative OSCE for pharmacy students in Qatar from multiple stakeholder perspectives.

Methods

Study Design

This was a qualitative research study using focus group methodology. Focus groups were chosen as the primary methodology for the ability to explore concepts in greater detail than what could be accomplished using survey or interview methodology (Merriam, 2014). This qualitative analysis supplemented a comprehensive psychometric analysis to gain further understanding regarding the strengths and limitations of this type of assessment.

Setting

The study was conducted in Doha, Qatar at Qatar University College of Pharmacy. The College received full accreditation status for the Bachelor of Science in Pharmacy degree (entry-to-practice degree) in 2012 from the Canadian Council for Accreditation of Pharmacy Programs and is therefore striving to implement educational initiatives in line with Canadian standards. Qatar is a small affluent country bordering Saudi Arabia and the Arabic Gulf. The population is diverse, with local Qatari comprising only approximately 20 percent (Permanent Population Committee, 2012). Expatriates comprise the majority of the population and come from all world regions, primarily South Asia, East Asia, Middle East, North Africa, and Western countries (Mahgoub & Qawasmeh, 2012). Healthcare and education sectors display similar ethnic diversity of working professionals. As part of a national vision, these sectors are undergoing major reforms to be in line with North American and European models and standards (General Secretariat for Development Planning, 2008).

OSCE Implementation

The 2015 (second iteration) OSCE consisted of ten interactive stations, each requiring students to interact with standardised actors. Two pharmacist assessors (faculty or clinicians) independently assessed each station. Assessment tools consisted of an analytical checklist for content and a five-point Likert global assessment scale. Cases were blueprinted according to adapted educational outcomes and local practice considerations (AFPC, 2010). Case competencies included health professional communication, counselling, adverse drug reaction management, minor ailment management, patient self-care, dose optimisation, pharmacy management, referral, and provision of drug information. Ten case groups consisting of four - five faculty members and clinicians wrote the cases and associated analytical checklists according to standard templates. Each group subsequently validated a case written by a different group. Prior to the OSCE day, assessors and standardised actors were recruited and trained. Assessors originated from many world regions, including the Middle East, North America, sub-Saharan Africa, and South Asia. Standardised patients were largely from Western and Arab countries and were primary lay people, with the exception of six who had pharmacy training. Exam centre staff personnel (time keepers, runners, hall monitors, track coordinators, and registration personnel) were also recruited. The OSCE took place over two cycles on one track and students were secluded during times of potential interaction to avoid sharing of information between cycles. All exam materials and references were printed in hardcopy and students were identified on exam booklets using coded stickers that were passed to assessors upon entrance to the station room. Cut scores were determined using the borderline regression method once exam results were obtained. OSCE candidates consisted of 21 graduating students from the Bachelor of Science in Pharmacy programme and five students entering the internship phase of the part-time Doctor of Pharmacy program.

Study Procedures

Purposive sampling was used to recruit candidates, standardised actors, assessors (both faculty and clinicians), and exam centre staff for participation in this study. Purposive sampling was used to ensure a variety of perspectives would be given, specifically pertaining to student performance levels. Original contact was made through email. Once interest was expressed for participation, consent and confidentiality agreements were obtained and the participant was scheduled for one of two pre-planned focus groups.

Participants arrived to the focus group and were greeted by facilitators. Facilitators were not involved in the design, implementation, or analysis of the OSCE to ensure participants felt comfortable to speak freely. One facilitator had previous experience with interviews and focus groups and took the lead during the discussion, while the other facilitator documented field notes and topics using a whiteboard and markers. The focus groups followed a pre-defined topic guide consisting of four major domains. Specifically, participants were to be asked their perspectives on the strengths, weaknesses, opportunities, and challenges associated with the 2015 OSCE. Each focus group was planned for one hour and was audio recorded. Facilitators managed discussion and time to ensure all topic domains were covered and all participants were encouraged to express their opinions equally. Photos of the points on the whiteboard were taken after the completion of each focus group.

Data Analysis

One investigator produced transcripts of the focus group verbatim. An additional investigator listened to the transcript and made corrections to the transcript, where applicable. During transcription, participants were

assigned a unique code (*i.e.* Student 1) to protect confidentiality of statements given. An inductive analytical approach was used for content analysis (Thomas, 2006). Transcripts were broken down into individual ideas or phrases that represented a single idea or concept. A coding framework was originally developed using whiteboard photos of documented points during the discussions. This framework was used to assign a code to each idea or phrase. Any idea or phrase that did not fit a particular code was given a new code and any point not supported by the transcripts was removed. Transcripts were read over multiple iterations until a final stable coding framework was established and major categories identified according to an iterative framework for qualitative data analysis (Srivastava & Hopwood, 2009). A second investigator independently reviewed the coded transcripts to check the original coding. Any identified discrepancies were resolved through discussion between the original coding investigator and the second investigator.

Data Trustworthiness

Numerous measures were taken to ensure trustworthiness of data. First, the faculty members and staff responsible for implementation of the OSCE were not facilitators or invited to the focus group. Second, two facilitators were present during the focus groups to ensure facilitation occurred without personal biases or interference. Thirdly, a second, independent investigator validated all the transcription and coding processes. Fourth, identified themes were supported with direct quotes arising from the transcripts.

Results

Participants

The first focus group consisted of two assessors from Qatar University, two students, one standardised actor, and two exam centre staff. The second focus group consisted of two assessors from Qatar University, one assessor from practice, one student, two standardised actors, and one exam centre staff. A total of 103 minutes were recorded over both focus groups that resulted in 37 pages of transcription. After conduction of two focus groups, it was deemed saturation was reached and no further focus groups were planned or conducted.

Main themes and codes

The focus group discussion resulted in 20 main categories that are distributed in the four main themes as shown in Table I. Overall strengths included training, assessment practices, familiarity, standardisation, and satisfaction. The weaknesses included discomfort, assessment practices, exam organisation, and training. The opportunities consisted of future licensure, regulator buy in, improvement, and standardised actor pool. The challenges included novelty, failure policies, specialised pharmacists, preparation of practicing pharmacists, collaboration, cultural differences, and OSCE overall scoring.

Supporting evidence

Strengths

Training was a common strength identified by assessors, standardised actors, and students. Assessors believed that they received a strong training from experts in the field:

> *"I think one of the strengths of it is that we've actually got through the programme. We've actually had formal training from somebody who has had experience in this area, so I feel like got the process better than a lot of schools."* (Internal Assessor 1)

Standardised patients acknowledged they received adequate training. From a different perspective, students felt formative OSCE exams embedded within the undergraduate programme helped prepare them for this cumulative version:

> *"I think of strength… the fact that we had a mock OSCE; we understood the actual set up."* (Student 2)

Many assessment practices associated with the OSCE were well received. For instance, students felt that the grading system, consisting of both analytical and global components, was beneficial for them. Both students and assessors commented that the use of multiple assessors per station was a strength. Specifically, students believed this aspect provided balance in grading:

> *"When I think of strength… There was more than one assessor. Because not everyone marks the same, so it felt like balance."* (Student 2)

Standardised patients noted that the pharmacy students were given enough time to read and understand their role before entering to their stations. Finally, assessors believed the collaborative nature in assessment between faculty members and practicing pharmacists was a positive design of the exam.

It also appeared from pharmacy students' responses that familiarity was important for their performance in the OSCE. Familiarity ranged from recognising some assessors and exam staff, as well as being familiar with the references provided for use:

> *"We know what resources we are going to use because Dr. "X" posted the name of the books, so we understood that, so it wasn't a total surprise."* (Student 1)

Interestingly, a standardised actor perceived as being the only one assigned to his or her station was seen as a strength of the exam in terms of standardisation and consistency:

> *"It was better. It was kind of consistent. I knew what to expect, what to say, and I think also that the assessors said it was good."* (Standardised patient 1)

Finally, student satisfaction was generally positive and they felt the exam was not stressful in nature:

> *"I think it [the exam] wasn't stressful."* (Student 1)

Table I: Final coding framework outlining identified strengths, weaknesses, opportunities, and challenges

Strengths:

1. Training
    - 1.1. Assessors
    - 1.2. SPs
    - 1.3. Students
        - 1.3.1. Mock OSCE
        - 1.3.2. Professional Skills Courses
2. Assessment
    - 2.1. Mutual grading system (analytical and global scoring)
    - 2.2. Skills diversity
    - 2.3. Multiple assessors
    - 2.4. Practice resemblance
    - 2.5. Time
    - 2.6. Collaboration
        - 2.6.1. Case building
        - 2.6.2. Assessment
3. Familiarity
    - 3.1. Assessors
    - 3.2 Resources
4. Standardization
    - 4.1. SP consistency
5. Satisfaction:
    - 5.1. Students

Weaknesses:

1. Discomfort
    - 1.1. Assessors
        - 1.1.1. Refreshments
        - 1.1.2. Rest
        - 1.1.3. Exam duration
    - 1.2. Students
        - 1.2.1. Refreshments
        - 1.2.2. Assessor unprofessionalism
2. Assessment
    - 2.1. Standardization
        - 2.1.1. Door instructions
        - 2.1.2. Resources feasibility
        - 2.1.3. Standardized patients
        - 2.1.4. Case validation
        - 2.1.5. Practice resemblance
    - 2.2. Grading
        - 2.2.1. Fairness
        - 2.2.2. Subjectivity
        - 2.2.3. Checkmark system
    - 2.3. Time
        - 2.3.1. Students
            - 2.3.1.1. Sticker experience
            - 2.3.1.2. Student readiness
        - 2.3.2. SPs
            - 2.3.2.1. Case familiarity
        - 2.3.3. Assessors
            - 2.3.3.1. Case familiarity
            - 2.3.3.2. Reflecting in global assessment
3. Exam Organization
    - 3.1. Insufficiency
        - 3.1.1. Assessors
        - 3.1.2. SPs
        - 3.1.3. Space
    - 3.2. Lack of coordination
    - 3.3. Interaction
        - 3.3.1. Students
        - 3.3.2. SPs
        - 3.3.3. Other personnel
4. Training
    - 4.1. Students
    - 4.2. Assessors
        - 4.2.1. Cultural communication assessment

Opportunities:

1. Future licensure
2. SCH buy in
3. Improvement
    - 3.1.Grading system
        - 3.1.1. SP involvement
    - 3.2. Bell system
    - 3.3. Sticker system
    - 3.4. Training
        - 3.4.1. SP
            - 3.4.1.1. Customized script
            - 3.4.1.2. Instructions
        - 3.4.2. Students
        - 3.4.3. Assessors
            - 3.4.3.1. Cultural communication assessment
            - 3.4.3.2. SP roles
    - 3.5. Exam resources
    - 3.6. Curriculum
    - 3.7. Recruitment
    - 3.8. Exam timing
    - 3.9. Door instructions
4. SP Pool

Challenges:

1. Novelty of OSCE idea in Gulf
2. Dealing with fails
3. Specialized pharmacists
4. Preparation of practicing pharmacists
5. Collaboration
6. Cultural difference
7. OSCE overall scoring

Weaknesses

For weaknesses, both students and assessors mentioned that the physical environment of the examination was uncomfortable at times. Some reasons identified were the need for refreshments because the exam duration was long. In addition, assessors noted that not having a scheduled break during the exam cycle was a deficiency. From a student perspective, they felt uncomfortable in some situations based on assessor behaviour:

> *"… and I think the assessor is bored or something. I was talking. He was sitting on the side, he was bored. He was like (knocking on the table)" "…or someone smiles, we know that we said something stupid."* (Student 1)

With respect to assessment practices, concerns were noted regarding standardised actor performance. Students commented on consistency of actor performance, while assessors commented on the quality and type of information given to students:

> *"For the standardised patients, some of them might, I know it's hard to be consistent, but some of them will give like different information or extra information or something."* (Student 1)

> *"Sometimes they would offer too much more than what was needed ... maybe because they were either pharmacists or healthcare providers. They weren't really playing the role of the patient."* (Internal assessor 1)

For validity relating to assessment practices, assessors thought that not using standardised actors in case validation was a weakness, while students focused more on how the cases reflected practice in Qatar itself:

> *"We are in Qatar, we went on rotation, some of the recommendations that they usually make is not similar to what we had to do here."* (Student 2)

Weaknesses related to grading resulted in three sub-themes: fairness, subjectivity, and checkmark system. Assessors commented that focusing on the analytical checklist distracted from the ability to assess global skills:

> *"As an assessor on one of the cases, I was focused a lot on the checklist, analytical checklist, just trying to make sure that, you know, I didn't miss something, you know that student said that I have to tick it. So so much focus there and then at the end you go to the global assessment, which is communication and you kind of feel, you know, I wasn't really focused on that, that much, and I don't have much time. The runner is coming in to take the sheet, so I think we had to do this so quickly and we didn't have much time to reflect."* (Internal assessor 2)

Other grading weaknesses included perceptions from students that assessors were poor communicators themselves and assessors believed checking off points in front of students may have been distracting and was taking away from their performance.

Under the theme of assessment practices, "time" emerged as a major sub-theme. Students reflected on time considerations within the station itself, where they had to give coded stickers to the assessors to place on their grading sheets before reading instructions and interacting with standardised actors. Assessors also reflected on time but in the context of requiring more time to better evaluate global performance. Finally, standardised actors and assessors believed that the actors needed more time to be familiar with their cases before the exam:

> *"But I think we need some more time for practice, to get to know the case."* (Standardised patient 1)

> *"Some more time should be spent in orienting the SPs, between the assessors and the SPs so that they know the case very well to know how to answer."* (Internal assessor 3)

Organisational factors such as space and logistics were also deemed weaknesses. For example, students believed there was a lack of coordination between all the stations during the exam due to inconsistencies using both a bell and knock system to alert when the station was to finish:

> *"I think there was a lack of coordination between the two sides, between the different sections, like when they ring the bell, then they will knock a bit later."* (Student 2)

Training emerged as a sub-theme for weaknesses. In this context, students believed adequate training was not provided to them to be successful on stations addressing competencies outside of a traditional patient care setting (*i.e.* management, advocacy). As well, assessors believed more training should be provided regarding cultural considerations, especially for communication practices within the local context.

Opportunities

Many opportunities emerged such as the possibility of using the OSCE for future licensure, the support of the regulator body, and the shared pool of standardised actors between different colleges or programmes. Also, it was mentioned that some weaknesses could be turned into opportunities if improved. Some of the participants added solutions such as using standardised patients as assessors of communication skills or using within-station bells to solve the problem of the bell system. Others proposed creating customised scripts for standardised patients to decrease their confusion, providing greater student training on competencies aside from patient care, conducting cultural communication training, training individuals to do both roles of assessor and standardised patients to solve recruitment problems, and providing multiple resources in every station such as electronic tablets.

> *"I think they [standardised actor] will probably do a better job than the assessors. Yeah, because, you know, you're communicating to me, it's about how I felt now, not how someone was thinking how I felt through this communication."* (Internal assessor 2)

> *"We will have like a communication course, like two hours, three hours, you know, [cultural] communications course. We are all on the same page, this is good communication, and this is bad communication. This is ideal, this is not."* (Standardised patient 3)

Challenges

Challenges identified consisted of many points related to both current assessment practices, as well as potential for future implementation as a national exam. Challenges related to the current state of the exam included dealing with failures, addressing cultural considerations with respect to communication, and enhancing collaboration between the college and practicing clinicians in exam design:

> *"Because it's like a high stake exam, if students fail, what would we do? Do we run another OSCE for them?"* (Internal assessor 1)

> *"I could watch that communication and say they didn't show respect. For you, it was fine; from my culture point of view, it was good, they could impress*

them…whatever…There is a culture issue here." (Internal Assessor 2)

In terms of challenges related to future adoption of the OSCE as a national licensing exam, assessors felt a lot of training would be required for practicing pharmacists to succeed and that there may be unfair obstacles for specialised pharmacists to pass a non-specialised, general exam.

"If this is something that we will roll out to all the practicing pharmacists in general, honestly, I don't think they will pass our OSCE. They need a lot of preparation prior to this be even used for a licensing exam." (Internal assessor 1)

Discussion

This study used qualitative methodology to critically analyse a cumulative OSCE for pharmacy students adapted into a Middle Eastern exit-from-degree context. The use of multiple stakeholders (students, assessors, standardised actors, and exam centre staff) enabled exploration of many strengths, weaknesses, opportunities, and challenges from varying perspectives. These considerations can be used to generate recommendations for improvement of the OSCE for future cycles. Critical discussion of our findings follows below.

The process of adaptation of a simulation assessment into a new context requires refinements to avoid different types of biases that could be associated with the adaptation process (Hambleton *et al*., 2004). A typical process of adaptation is lengthy and requires certain procedures, one of which is performing validation analyses (Geisinger, 1994). The qualitative analysis detailed in this paper provides validity data regarding the assessment itself, as well as its content. An integrated analysis of identified strengths, weaknesses, opportunities, and challenges can identify gaps and expose limitations of current assessment practices, as well as provide quality improvement data for future assessment iterations.

A major strength of the exam was perceived to be the use of multiple assessors per station. It was deemed this was an important consideration to achieve balance across assessors and provide fairness in grading. This is in line with current beliefs in assessment, where the use of multiple assessors can provide differing yet valid interpretations of performance (Ginsburg *et al*., 2010). This concept is particularly relevant for the global assessment and evaluation of communication skills. In our multicultural setting, it is possible that individual assessors perceive communication behaviours differently and therefore balance can be obtained by using multiple assessors per station. We recommend that this practice continue in our setting, as well as other settings with high cultural diversity.

Many weaknesses were identified and a few important ones should be discussed. Concerns related to content validity emerged from the analysis. It was signalled that some stations may not have reflected current practice in Qatar and/or focused on competencies not addressed within the undergraduate curriculum. Specifically, students mentioned a station requiring them to provide remedial feedback for a pharmacy technician. It is possible that this identified a learning gap in the curriculum and/or experiential training activities and those students should be expected to be competent with these skills upon graduation. Therefore, we recommend assessment of the curriculum and practice site activities to determine how to address this identified need. Also, we recommend future cycles to include stations blueprinted to competencies aside from "Care Provider" and "Communicator", in order to provide more opportunities for identifying curricular gaps and learning needs.

Students identified concerns with content of another station related to reflection of practice in Qatar. This station required students to assess a patient presenting with a prescription to a community pharmacy for azithromycin and determine that he needed referral back to the physician due to risk factors and symptoms specific for tuberculosis. Students felt this was outside of the pharmacist's scope of practice in Qatar. While community pharmacy practice is largely underdeveloped in Qatar, patient assessment is a core competency expected of graduates. However, the problem may have been the setting of the case, as this patient likely would have received a prescription from a hospital or clinic. Therefore, we recommend a focus on setting and problem alignment during case validation procedures at the time of case writing.

Opportunities focused on solutions to overcome weaknesses, however the potential for the exam to be adapted as a national licensing assessment was interesting. While this approach would enable benchmarking practitioner competencies against predefined performance levels, many considerations must be addressed before this could occur. Some of these considerations include the validity of an English-language assessment in a multilingual society, determination of passing standards for both local and foreign trained pharmacists, and feasibility in terms of personnel and resources required to implement. Despite these potential barriers, future studies could assess the use of OSCE as an entry-to-practice or continual competency assessment within Qatar or the greater Gulf region.

Two major challenges were identified, which were failure policies and cultural considerations with assessments. Currently, the OSCE is not a 'must pass' exam in our setting due to its novelty and uncertain validity and reliability as a high stakes assessment. However, if future iterations can demonstrate acceptable psychometric properties, it is fair to consider moving it towards a mandatory pass prior to graduation. The purpose of this would be to ensure graduates meet predefined minimal competency expectations prior to entering practice. Regarding cultural considerations in assessment, this is a very interesting finding that relates to most settings in the world but specifically those with high cultural diversity. It

is possible assessment of communication skills, professionalism, and other competencies may be influenced by an assessor's culture. Future studies are warranted to determine the nature and extent of this issue and perhaps incorporate qualitative assessment practices to better capture these differing perspectives of performance.

The points discussed above explore strengths, weaknesses, opportunities, and challenges of the OSCE yet some limitations should be noted. First, the analysis was limited to a single iteration of a cumulative OSCE conducted in one setting. While this decreases generalisability, other centres can learn from our findings especially in terms of considerations such as exam set-up (*i.e.* assessor conditions), grading decisions, cultural considerations, and assessment of content validity. Second, the identified perceptions do not likely reflect the entire population sampled, especially for student participants. However, points were deemed similar between both focus groups and therefore it is quite possible most major ideas were documented using our methodology. Lastly, the exploratory nature of the study design resulted in many different categories and themes to address. As such, the discussion attempted to focus on key points that could be related to other settings and contexts.

Conclusion

This study identified many strengths, weaknesses, opportunities, and challenges regarding the quality, utility, and value of a cumulative OSCE in a Gulf context using a novel focus group approach analysing key stakeholder perceptions. Findings resulted in many recommendations for refinement for future cycles, in order to improve validity of the exam as a high stakes assessment context. Specifically, assessor conditions and cultural considerations in assessment emerged as topics to be further explored in future studies. Identified perceptions and associated recommendations should be considered by other centres considering implementing a similar OSCE in their settings.

Financial Disclosure Summary

No funding was provided for this project. All authors report no conflicts of interest.

References

Abdelaziz, A., Hany, M., Atwa, H., Talaat, W. & Hosny, S. (2016). Development, implementation, and evaluation of an integrated multidisciplinary Objective Structured Clinical Examination (OSCE) in primary health care settings within limited resources. *Medical Teacher*, **38**, 272-279.

AFPC. (2010). Educational outcomes for first professional degree programs in pharmacy (entry-to-practice pharmacy programs) in Canada (online). Available at: https://www.afpc.info/sites/default/files/AFPC%20Educational%20Outcomes.pdf. Accessed 29th June, 2016.

Al-Azzawi, A.M.J., Nagavi, B.G., Hachim, M.Y. & Mossa, O.H. (2016). The implementation and development of an objective structured clinical examination in the community pharmacy course of a select Gulf-region academic institution (Ras Al Khaimah College of Pharmaceutical Sciences): a pilot study. *Innovations in Education and Teaching International*, **53**, 60-72.

Geisinger, K.F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. Psychological Assessment, 6, 304-312.

General Secretariat for Development Planning. (2008). Qatar National vision 2030 (online). Available at: http://www.gsdp.gov.qa/wwwl_docs/QNV2030_English_v2.pdf. Accessed 29th June, 2016.

Ginsburg, S., McIlroy, J., Oulanova, O., Eva, K. & Regehr, G. (2010). Toward authentic clinical evaluation: pitfalls in the pursuit of competency. *Academic Medicine*, **85**, 780-786.

Gronlund, N.E. (1998). Assessment of student achievement (6th edition). Boston MA: Allyn & Bacon.

Hambleton, R.K., Merenda, P.F. & Spielberger, C.D. (2005). Adapting educational and psychological tests for cross-cultural assessment. Mahwah NJ: Lawrence Erlbaum.

Johnson, B., Pyburn, R., Bolan, C., Byrne, C., Jewesson, P., Robertson-Malt, S., El-Tawil, M. & Verjee, M. (2011). Qatar Interprofessional Health Council: IPE for Qatar. *Avicenna*, **2**.

Karim, J.A., Marwan, Y.A., Dawas, A.M. & Akhtar, S. (2012). Self-confidence of medical students in performing clinical skills acquired during their surgical rotation. Assessing clinical skills education in Kuwait. *Saudi Medical Journal*, **33**, 1310-1316.

Mahgoub, Y. & Qawasmeh, R.A. (2012). Cultural and Economic Influences on Multicultural Cities: The Case of Doha, Qatar. *Open House International*, **37**, 33-41.

McMillan, J.H., Hellsten, L. & Klinger, D. (2007). Classroom assessment: Principles and practice for effective standards-based instruction. Boston, MA: Pearson/Allyn & Bacon.

Merriam, S.B. & Tisdell, E.J. (2015). Qualitative research: A guide to design and implementation. San Francisco CA: Jossey-Bass.

Permanent Population Committee. (2012). Qatar Population Statistics 2012: Three Years After Launching the Population Policy (online). Available at: http://www.gsdp.gov.qa/portal/page/portal/ppc/PPC_home/ppc_news/ppc_files_upload/populations_status_2012_en.pdf. Accessed 29th June,2016.

Raheel, H. & Naeem, N. (2013). Assessing the Objective Structured Clinical Examination: Saudi family medicine undergraduate medical students' perceptions of the tool. *Journal of Pakistan Medical Association*, **63**, 1281-1284.

Selim, A.A., Ramadan, F.H., El-Gueneidy, M.M. & Gaafer, M.M. (2012). Using Objective Structured Clinical Examination (OSCE) in undergraduate psychiatric nursing education: Is it reliable and valid? *Nurse Education Today*, **32**(3), 283-288.

Srivastava, P. & Hopwood, N. (2009). A practical iterative framework for qualitative data analysis. *International Journal of Qualitative Methods*, **8**(1), 76-84.

Swanson, D.B., Norman, G.R. & Linn, R.L. (1995). Performance-based assessment: Lessons from the health professions. *Educational Researcher*, **24**(5), 5-11.

Thomas, D.R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, **27**(2), 237-246.

Turan, S. & Konan, A. (2012). Self-regulated learning strategies used in surgical clerkship and the relationship with clinical achievement. *Journal of Surgical Education*, **69**(2), 218-225.

Wiggins, G. (1989). A true test. *Phi Delta Kappan*, **70**(9), 703-713.

Wiggins, G.P. (1993). Assessing student performance: Exploring the purpose and limits of testing. San Francisco CA: Jossey-Bass.

Zahid, M.A., Al-Zayed, A., Ohaeri, J. & Varghese, R. (2011). Introducing the objective structured clinical examination (OSCE) in the undergraduate psychiatric curriculum: evaluation after one year. *Academic Psychiatry*, **35**(6), 365-369.