Ensuring robust OSCE assessments: a reflective account from a Scottish School of Nursing.

Abstract:

AIM: This paper reflects on the experience of one Scottish University in conducting a face-to-face Objective Structured Examination (OSCE) for large cohorts of student nurses. It outlines the challenges experienced and learning gained.

Borton's model of reflection frames this work due to its simplicity, ease of application and cyclical nature.

Background: The theoretical framework for the OSCE is critical thinking, enabling students to apply those skills authentically. OSCE's are designed to transfer classroom knowledge to clinical practice and offer an authentic work-based assessment.

Design: Validity and robustness are key considerations in any assessment, and in OSCE, the number of stations that students encounter is important and debated. We used a case-study based OSCE approach initially over four stations and following reflection, changed to one long station with four phases.

Results: In OSCE examinations, interrater reliability is a necessity, and students expect equity of approach. We identified that despite clear marking criteria, marks were polarised, with students achieving high or low marks with little middle ground. Review of examination papers highlighted that although students' overall performance was good some had failed in at least one station, suggesting a four-station approach may skew results. On reflection we hypothesised that using a one station case study-based, phased approach enabled the examiner to build up a more holistic picture of student knowledge and skills. It also provided the student opportunity to develop a rapport with the examiner and standardised patient, thereby putting them more at ease. We argue that this approach is holistic, authentic and student centred.

Conclusions: Our experience highlights that a single station, four phase OSCE is preferrable, enabling students to integrate all aspects of the assessment and provides a holistic view of clinical skills and knowledge.

Key Words

Objective structured clinical examination; student nurses; assessment; validity; reflection

Background

The first Objective Structured Clinical Examinations (OSCE) were developed by Harden et al. (1975), to assess the clinical competency of medical students and provide a structured, controlled approach to assessment. Since then, OSCE's have adapted for use in nursing and allied health professions as they provide an opportunity to standardise assessments and create equal conditions for all students (Lyngå et al., 2019). OSCE's are a form of work-based assessment that lends authenticity to the assessment process and both students and assessors consider them to be a valid, reliable method for assessing clinical skills and knowledge application (Lyngå et al., 2019; Mitchell et al., 2009). Educators consider them an

effective means to assess the application and synthesis of classroom knowledge to clinical practice (Aryal et al., 2021; Zhu et al., 2017), and they have both formative and summative properties by promoting continuous learning (Torsney et al., 2014).

We used the OSCE exam in two settings; in the first module of students' second year of an undergraduate nursing programme, and in a top-up degree for registered nurses delivered in Singapore. As such neither exam is high stakes. High stakes OSCE's typically assess students at the end of their programme and determine overall course success (Blythe et al., 2021; Shulruf et al., 2018). Interestingly, although we assessed two quite different cohorts of students but there was minimal difference in overall performance with a fail rate of 15% in the qualified cohort and 18% in the undergraduate group.

Learning Theory

The theoretical framework for the OSCE is critical thinking (Zhu et al., 2017), enabling students to apply problem solving skills in an authentic context. Within the OSCE, we assessed both knowledge and behaviour. According to Khattab and Rawlings (2001), knowing how and knowing why are equally important and both should be assessed.

Dehaene (2021) theory of learning provides a useful lens to evaluate the process underpinning learning from the OSCE. It also provides a useful theory to explain why OSCEs should work in the first place. An OSCE is an 'authentic assessment'. Authentic assessments are assessments that pertain to the real world of whatever is being assessed and have been used for decades to support learning (Villarroel et al., 2020). As introduced above, OSCE is an example of an authentic assessment because it is 'work based' learning. The idea is to mimic practice and test skills likely to be useful in real world practice. According to Sokhanvar et al. (2021) authentic assessment can improve student satisfaction, increase goal-oriented effort as well as enhancing engagement more generally. It facilitates autonomy, self-regulation, and metacognition (Villarroel et al., 2020).

Reflection and Reflective Practice

The concept and theory of reflection and reflective practice has a long history. The philosopher and educator John Dewey described reflection as "The active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends" (Dewey, 1933) p118).

Reflective practice is fundamental to the integration of nursing theory and practice (Barbagallo, 2021). It allows nurses to interpret and analyse their experiences and, by using a reflective model, produces an actionable outcome, leading to personal or professional development and improve the quality of care (Galutira & Domingo, 2018).

Borton's Model of Reflection

Borton's model of reflection frames our reflections on the OSCE due to its simplicity, ease of application and its cyclical nature. Borton's model proposes three questions: "What?", "So what?" and "Now what?". The first question asks us to consider what has happened, the second is to guide us to try and make sense of what has happened and evaluate events and the final to consider a way forward and our responses should the event happen again (Jasper et al., 2013; Smith et al., 2016)).

Reflection on OSCE Format

What?

We used a case-study based OSCE to assess six hundred (600) student nurses in the first trimester of their 2nd year and over two hundred (200) registered nurses undertaking a 'top-up' degree in Singapore. In Scotland, OSCE's took place over ten days in a purpose-built clinical skills and simulation centre. Over forty academic staff and twenty simulated patients were involved highlighting the significant resources required. In Singapore, a team of eight 'flying faculty' travelled to conduct the assessments. Faculty worked in pairs with one playing the role of patient and one being the assessor. Examinations were conducted in either the small 'nursing lab' and classrooms. All OSCEs in Singapore were recorded to facilitate external moderation and once moderated were deleted. These recordings also aided this reflection.

In both modules, the teaching concentrated on patient assessment and clinical decision making within case studies that followed a patient journey from acute admission to discharge. Learning outcomes are as follows:

- LO1: Demonstrate how to undertake a comprehensive, systematic assessment of patients' physical health.
- LO2: Demonstrate how to undertake an assessment of patients' general health and wellbeing.
- LO3: Interpret data gained the patients' health assessment to enable an evidence-based plan of care to be formulated.
- LO4: Explore, identify, and evaluate appropriate evidence and use it to formulate a plan of care.

Preparation for OSCE is an important consideration and it is already known that the ability to 'rehearse' improves student's overall performance (Al Rushood & Al-Eisa, 2020). We prepared students for examination through delivery of theory, case study-based application, and clinical skill classes with an opportunity for 'mock' OSCE.

Within the OSCE, there were essential criteria that students needed to meet to pass, and details of these were available to students. Initially the OSCE comprised four eight-minute stations, examining students in key clinical skills in the context of a specific case study in: adult health, mental health, child health or learning disabilities. Examiners presented students with a handover, and students moved from room to room to complete individual aspects of the patient's journey from history taking, systematic assessment, identification of

care priorities, through to medication advice and review of best practice guidelines. Movement between these phases, or 'stages' was prompted by a timed bell. Once students had moved on there was no further opportunity to add to the responses provided in that phase. Student feedback highlighted that the call bell increased their anxiety, and, on moderation, results revealed there was inconsistency in student performance across stations.

So what?

We noted that despite clear marking criteria, marks were polarised, with students achieving extremely high or extremely low marks with little middle ground. Review of papers highlighted those students who failed had performed badly in at least one stage, but their overall performance was generally good. Consequently, the team agreed to trial a continuous assessment. This included introducing a four-phase approach with the student remaining in the same room with the same examiner and simulated patient. Students were permitted to return to previous phases of the examination and add or correct their responses. An example of this was where students forgot to enquire about patient allergies in phase 1 patient history but recalled this in phase 4 when discussing medications.

We evaluated the benefits of four discreet stations, versus overall phased approach, and tried both. Feedback from key stakeholders, including students, examiners and external moderators endorsed the one station, four phase approach, noting that students were more able to develop a rapport with the examiner and standardised patient. We argue that this approach, assesses examinees overall performance and facilitates holistic student–centred assessment (Harden, 2015).

Of course, in making this change the key concern raised was academic rigour. Validity and robustness are key considerations in any assessment, and in OSCE, the number of stations that students encounter is important and debated. Some theorists claim that many stations are required to ensure rigour (Mitchell et al., 2009) as this allows several examiners to judge the student's performance and allows students more opportunities to succeed. This approach is consistent with the original OSCE that assessed medical students in a range of discrete clinical skills over 16- 20 stations. A criticism of this approach is that the wider context of clinical practice and holistic aspects of working with people are not examined (Lyngå et al., 2019). Crossley, (2012) argues that objective tick boxes can deconstruct the assessment and fail to examine the underpinning knowledge and understanding necessary for competence. For this reason, Mitchell et al. (2009) argue that a one station OSCE is preferrable to integrate all aspects of the assessment and provide a holistic view of clinical skills and knowledge. Indeed, since their inception OSCEs have evolved and comprise fewer stations and focus more on total client consultation or patient journey (Mitchell et al., 2009; Smith et al., 2012).

Now What?

As we have already identified the purpose of the OSCE is to transfer classroom knowledge to clinical practice (Zhu et al., 2017) and offer an authentic work-based assessment that promotes continuous learning (Torsney et al., 2014). We have shown that OSCE assessment

is a valid tool for this purpose. However, care must be taken in considering number and format of stations. Yuan (2021) asserts that a case study model is holistic, authentic and enables transfer of theoretical knowledge into clinical practice. In our experience an OSCE, structured to follow the patient's journey, from acute admission to discharge and ongoing care, promoted an integrated approach and produced more consistent student outcomes. Furthermore, the resource implications were positive, requiring far fewer staff, a key consideration within the contemporary context of tertiary education. As a result, we have embedded this approach in our home and Singapore programmes.

Reflection on OSCE Scoring

What?

We produced scoring papers that provided objective criteria on which students' performance was evaluated (see appendix 1 for example). Examiners included members of the module teaching team and other volunteer faculty members. All markers attended a training session and had copies of the OSCE documentation and expected responses. Standardised patients participated in two phases involving history taking and providing patient discharge medication advice. The systematic assessment component was performed on mannequins. Pre-briefing instructions and scripts were provided for examiners and simulated patients, alongside local preparation for this role. Despite this, students reported variability in the approach and demeanour of both assessor and standardised patient that affected their performance.

Interrater reliability is a necessity in OSCE examinations and students expect equity of approach (Govaerts & van der Vleuten, 2013; Kogan et al., 2015). We identified that despite clear marking criteria, there was ambiguity around student attainment. Anecdotally, some examiners prompted students whilst others were less supportive in their approach. For example, we asked students to discuss assessment findings in professional language, using correct terminology with the proviso to accept a correct answer colloquially stated. Examiners differed in the application of this, with some examiners stating that lack of professional language was synonymous with lack of knowledge, and that the overall impression the student had given was one of unpreparedness.

So what?

Another key challenge previously highlighted was the polarity of results that provided a skewed picture of attainment. This problem has also been highlighted by Beck and Sitzman (2019) who suggest that the tick box assessment structure of an OSCE lacks sensitivity, causing students achieve at extremes end of the marking spectrum. Furthermore, they report that students who performed poorly in one or more stations tended to achieve a low score overall and our results are consistent with this.

It is recognised that OSCE's are heavily reliant on assessor behaviour (Pell et al., 2010) and we noted that despite robust training in place there were some inconsistencies in marker judgements. Even within objective rating scales, markers can disagree on the interpretation of what they have observed (Crossley & Jolly, 2012). These authors go onto highlight how

aspects such as poor communication skills or professionalism can influence the examiner's judgement and predetermine if a performance is satisfactory or not. East et al. (2014) evaluated the impact of assessor characteristics and perceptions on students' results and found that less than half the assessors taught on the course and that this affected their marking resulting in more fails. As we have stated, the original OSCE format was heavily reliant on a large number of assessors, not all taught on this module. Our evaluation revealed that having fewer OSCE stations required fewer examiners. Therefore, we could effectively manage the assessment timetable within the module team, and this improved our first-time pass rate and lent greater consistency to the process.

Now what?

We acknowledge that mechanical checklists can stifle the ability of the examiner to build up a holistic picture of student achievement. For this reason, the OSCE was redesigned to include fewer high-level criteria, a global rating scale, and a pass/fail grade based on overall performance (Homer & Russell, 2021). We believe this approach to be student centred and amenable to robust evaluation. Undoubtedly, meticulous examiner training improves inter-rater reliability (Lyngå et al., 2019) and some successful strategies include the provision of more detailed support material (Pell et al., 2010). Babar and Afzal (2021) also recommends having a 'dry run' to enable examiners to experience the OSCE and highlight any anomalies prior to the examination day. To this end the team are developing a training opportunity that encompasses simulation and feedback well in advance of the examination day.

Conclusions and Recommendations for future Practice.

The reflections discussed in this paper discuss a real-world application of OSCE assessments in large cohorts of student nurses. Our experience highlights the challenges of ensuring consistency of assessment, despite robust objective criteria and examiner preparation. We have learned that a potentially skewed picture of attainment, and polarised results might be avoided by using a longer one station, four phased, case study-based approach, which enables the examiner to build up a holistic picture of student knowledge and skills. Feedback from our staff and students highlighted that another benefit of this approach to OSCE is the ability of the student to develop a rapport with the examiner and standardised patient, thereby putting them more at ease.

For future OSCEs we recognise that examiner preparation is key and recording of OSCE's can enable the team to revisit and reflect on personal performance and plan for greater consistency going forward. Equally, in the future simulated patient scripts will contain more specific detail so that the 'patient' consistently conveys the pertinent information about the patient's condition.

Undoubtedly, our team, will continue to reflect on the OSCE, student performance, and ensuring authenticity and robustness in assessment. We remain committed to the holistic approach in OSCE that more closely mimics real world practice.

References

Al Rushood, M., & Al-Eisa, A. (2020). Factors predicting students' performance in the final pediatrics OSCE. *PLoS One*, *15*(9), e0236484-e0236484. https://doi.org/10.1371/journal.pone.0236484

Aryal, K., Hamed, M., & Currow, C. (2021). The usefulness of work-based assessments in higher surgical training: A systematic review. *International journal of surgery (London, England)*, *94*, 106127-106127. https://doi.org/10.1016/j.ijsu.2021.106127

Babar, S., & Afzal, A. (2021). The "new-normal" OSCE examination: Executing in the COVID-19 era. *Pakistan journal of medical sciences*, *37*(7), 2026-2028. https://doi.org/10.12669/pjms.37.7.4568

Barbagallo, M. S. (2021). Nursing students' perceptions and experiences of reflective practice: A qualitative meta-synthesis. *Teaching and learning in nursing*, *16*(1), 24-31. https://doi.org/10.1016/j.teln.2020.07.006

Beck, M. S., & Sitzman, K. (2019). Compelling Reasons for Using Digital Stories to Teach: A Descriptive Qualitative Study. *Teaching and learning in nursing*, *14*(4), 265-269. https://doi.org/10.1016/j.teln.2019.06.007

Blythe, J., Patel, N. S. A., Spiring, W., Easton, G., Evans, D., Meskevicius-Sadler, E., Noshib, H., & Gordon, H. (2021). Undertaking a high stakes virtual OSCE ("VOSCE") during Covid-19. *BMC medical education*, *21*(1), 221-221. https://doi.org/10.1186/s12909-021-02660-5

Crossley, J., & Jolly, B. (2012). Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people. *Medical Education*, *46*(1), 28-37. https://doi.org/10.1111/j.1365-2923.2011.04166.x

Dehaene, S. (2021). *How we learn : the new science of education and the brain*. Penguin Books.

Dewey, J. (1933). *How we think : a restatement of the relation of reflective thinking to the educative process*. Heath.

East, L., Peters, K., Halcomb, E., Raymond, D., & Salamonson, Y. (2014). Evaluating Objective Structured Clinical Assessment (OSCA) in undergraduate nursing. *Nurse Education in Practice*, *14*(5), 461-467. https://doi.org/10.1016/j.nepr.2014.03.005

Galutira, & Domingo, G. (2018). Theory of Reflective Practice in Nursing. *International Journal of Nursing Science*, *8*(3), 5. https://doi.org/10.5923/j.nursing.20180803.02

Govaerts, M., & van der Vleuten, C. P. M. (2013). Validity in work-based assessment: expanding our horizons. *Medical Education*, *47*(12), 1164-1174. https://doi.org/10.1111/medu.12289

Harden, R. M. (2015). Misconceptions and the OSCE. *Med Teach*, *37*(7), 608-610. https://doi.org/10.3109/0142159X.2015.1042443

Harden, R. M., Stevenson, M., Downie, W. W., & Wilson, G. M. (1975). Assessment of clinical competence using objective structured examination. *British Medical Journal*, *1*(5955), 447-451. https://doi.org/10.1136/bmj.1.5955.447

Homer, M., & Russell, J. (2021). Conjunctive standards in OSCEs: The why and the how of number of stations passed criteria. *Medical teacher*, *43*(4), 448-455. https://doi.org/10.1080/0142159X.2020.1856353

Jasper, M., Rosser, M., Mooney, G. P., & Jasper, M. (2013). *Professional development, reflection, and decision-making in nursing and healthcare* (2nd edition. ed.). Wiley-Blackwell.

Khattab, A. D., & Rawlings, B. (2001). Assessing nurse practitioner students using a modified objective structured clinical examination (OSCE). *Nurse Education Today*, *21*(7), 541-550. https://doi.org/10.1054/nedt.2001.0590

Kogan, J. R., Conforti, L. N., Bernabeo, E., Iobst, W., & Holmboe, E. (2015). How faculty members experience workplace-based assessment rater training: a qualitative study. *Medical Education*, *49*(7), 692-708. https://doi.org/10.1111/medu.12733

Lyngå, P., Masiello, I., Karlgren, K., & Joelsson-Alm, E. (2019). Experiences of using an OSCE protocol in clinical examinations of nursing students - A comparison of student and faculty assessments. *Nurse education in practice*, *35*, 130-134. https://doi.org/10.1016/j.nepr.2019.02.004

Mitchell, M. L., Henderson, A., Groves, M., Dalton, M., & Nulty, D. (2009). The objective structured clinical examination (OSCE): Optimising its value in the undergraduate nursing curriculum. *Nurse Education Today*, *29*(4), 398-404. https://doi.org/10.1016/j.nedt.2008.10.007

Pell, G., Fuller, R., Homer, M., & Roberts, T. (2010). How to measure the quality of the OSCE: A review of metrics - AMEE guide no. 49. *Medical teacher*, *32*(10), 802-811. https://doi.org/10.3109/0142159X.2010.507716

Shulruf, B., Damodaran, A., Jones, P., Kennedy, S., Mangos, G., O'Sullivan, A. J., Rhee, J., Taylor, S., Velan, G., & Harris, P. (2018). Enhancing the defensibility of examiners' marks in high stake OSCEs. *BMC medical education*, *18*(1), 10-10. https://doi.org/10.1186/s12909-017-1112-z

Smith, S., James, A., Brogan, A., Adamson, E., & Gentleman, M. (2016). Reflections about experiences of compassionate care from award winning undergraduate nurses – What, so what … now what? *Journal of Compassionate Health Care*, *3*(1). https://doi.org/10.1186/s40639-016-0023-x

Smith, V., Muldoon, K., & Biesty, L. (2012). The Objective Structured Clinical Examination (OSCE) as a strategy for assessing clinical competence in midwifery education in Ireland: A critical review. *Nurse Education in Practice*, *12*(5), 242-247. https://doi.org/10.1016/j.nepr.2012.04.012

Sokhanvar, Z., Salehi, K., & Sokhanvar, F. (2021). Advantages of authentic assessment for improving the learning experience and employability skills of higher education students: A systematic literature review. *Studies in Educational Evaluation*, *70*(101030). https://doi.org/https://doi.org/10.1016/J.STUEDUC.2021.101030

Torsney, K. M., Cocker, D. M., & Slesser, A. A. P. (2014). The Modern Surgeon and Competency Assessment: Are the Workplace-Based Assessments Evidence-Based? *World Journal Of Surgery*, *39*(3), 623-633. https://doi.org/10.1007/s00268-014-2875-6

Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C. (2020). Using principles of authentic assessment to redesign written examinations and tests. *Innovations in Education and Teaching International*, *57*, 38-49. https://doi.org/10.1080/14703297.2018.1564882

Yuan, H.-B. (2021). Evaluation of undergraduate students' nursing assessment and communication skills through an objective structured clinical examination within a high-fidelity simulation using a student-simulated patient. *Frontiers of Nursing*, *8*(2), 159-168. https://doi.org/10.2478/fon-2021-0018

Zhu, X., Yang, L., Lin, P., Lu, G., Xiao, N., Yang, S., & Sui, S. (2017). Assessing Nursing Students' Clinical Competencies Using a Problem-Focused Objective Structured Clinical Examination. *Western Journal of Nursing Research*, *39*(3), 388-399. https://doi.org/10.1177/0193945916667727