The reliability and validity of the OSCE as an assessment of capability within advanced practitioner curricula
Stirling, Kevin; Henderson, Colette

Published in:
British Journal of Nursing

DOI:

Publication date:
2021

Citation for published version (APA):
Commissioned ACP comment piece:
The reliability and validity of the Observed Structured Clinical Examination as an assessment of capability within Advanced Practitioner curricula.

Dr Kevin Stirling
Lecturer in Nursing
Clinical Skills and Simulation Coordinator (Pre-registration Programme)

And

Colette Henderson
Lecturer in Nursing
Programme Lead MSc Advanced Practice
and Chair Scottish Advanced Practice Educators Network (SAPEN)

University of Dundee
School of Nursing and Health Sciences
11 Airlie Place, Dundee DD1 4HJ
01382 385907
Advanced Practitioners (AP) are healthcare professionals educated at master’s level to manage entire episodes of care (Health Education England, 2017; Scottish Government, 2017). Advanced practice is defined by the ability to demonstrate capability across four pillars of practice: 1) Clinical practice, 2) Facilitating learning 3) Leadership and 4) Evidence, research, and development (Health Education England, 2017; Scottish Government, 2017). Exact terminology varies across the UK. Programmes of education for advanced practitioners are aligned with the four pillars to enable practitioners to demonstrate capability in all four domains.

It has become accepted practice within AP curricula that, capability is assessed using the Observed Structured Clinical Examination (OSCE). The OSCE was instituted in Dundee, Scotland by Harden and Gleeson (1975) and has gone on to become widely adopted within undergraduate and postgraduate curricula globally. The OSCE is structured using multiple stations that students rotate round, completing tasks specific to that station. As with workplace-based assessments, each station assesses capability in relation to a specific skill or procedure. The justification for the continued use of the OSCE within AP curricula is to allow educators to standardise the assessment of capability at set intervals within a programme of study and to support the translation of knowledge into practice.

Boursicot et al. (2014) stated that the design of the OSCE (whereby each station uses structured marking schedules that ensures consistent scoring by examiners) created an acceptable environment to assess capability. The authors argue that the inclusion of multiple stations (focusing on different examination, communication, or procedural
skills by speciality) resulted in a more reliable picture of a participant’s overall capability as performance was being reviewed by multiple examiners (thus bias is minimised).

Brannick et al. (2011) conducted a meta-analysis to determine the reliability of the OSCE. In the instance of this study, data were analysed across items (reviewing student performance against a set criteria) and across stations (reviewing the performance of the OSCE itself). The authors identified that Cronbach’s alpha (α) was the most commonly used measure of internal consistency, so this became the principal comparator across the 39 studies which met the inclusion criteria (8% of the studies analysed were from AP curricula). The authors’ reported α=0.78 across items and α=0.66 across stations (the mean α for OSCEs with less than 10 stations was 0.56 and 0.74 with greater than 10 stations). The authors identified that the construct OSCE stations made it more difficult to reliably assess communication skills rather than a specific examination or procedural skill. A better than average reliability was associated with a greater number of stations and a higher number of examiners per station, although it was acknowledged that this would create an examination process that was logistically challenging and resource intensive.

The Advanced Practice team at the University of Dundee, Scotland have held iterative discussions with educators (both in the United Kingdom and Europe) who deliver AP curricula. These discussions have identified three issues relating to the construct and utilisation of the OSCE as a means of assessing the capability of APs. Firstly, there is a notable variation in the number of OSCE stations used within AP curricula to assess capability (variation ranged from 3-12 stations) therefore the reliability of this assessment process must be questioned. Secondly, the assessment instrument used
within OSCE station is not sophisticated enough to discriminate performance reliably across domains (it is common practice that only a pass or fail judgement is given for each station). Thirdly, the construct of OSCE stations is designed to demonstrate capability in managing entire episodes of care, therefore a mixture of stations are utilised to assess the capability of APs in relation to examination, communication and procedural skills. This approach (especially in relation to communication skills) has been shown to not deliver an objective assessment of capability (Brannick et al. 2011).

The COVID-19 pandemic necessitated a move towards remote assessment which meant that OSCEs were delivered within contexts where this assessment was not traditionally undertaken. This rapid adoption of a dispersed manner of assessment along with a national drive to instigate compassionate assessment practices (Quality Assurance Agency, 2020) challenged the Advanced Practice team to examine the validity of the OSCE as a mean of assessing the capability of APs within a programme of study and whether this assessment supports the translation of knowledge into practice. The authors are keen that the data collected from this survey is representative of current practice and the lived experience of APs so that the subsequent recommendations made are relevant to the current and future requirement of the profession and AP curricula.

Readers are invited to engage with this research study and anyone interested in doing so can contact Dr Kevin Stirling K.J.Stirling@dundee.ac.uk
Reference List


