

Cite as: MacNeily A. Hawks and doves, differing content, scoring systems, and delivery platforms: No wonder there's variability in grades. *Can Urol Assoc J* 2024;18(4):120. <http://dx.doi.org/10.5489/cuaj.8747>

See related article on page 116

Hawks and doves, differing content, scoring systems, and delivery platforms: No wonder there's variability in grades

In 1990, the Royal College of Physicians and Surgeons of Canada, in conjunction with the urology test committee, changed the format of the certification examinations, which for decades had been a two-part, multiple-choice exam, followed several months later by three in-person, one-hour, oral examinations. The new format was a short-answer component, a two-part multiple-choice exam, followed the next day with an Objective Structured Clinical Examination (OSCE). The rationale was that the OSCE was felt to be more objective and structured (hence the name) than an examiner using a case of some arcane urologic topic as a stepping point to interrogate the depths of a candidate's knowledge.

Although the timing of the RCPSC exams and the written component formatting have evolved, the OSCE has remained a constant. The QUEST program, initially a labor of love attempting to help Canadian trainees prepare for this new format, has also evolved into an annual rite of passage, and an opportunity for a dress rehearsal prior to the real quiz. The OSCE also remains at the core of QUEST.

In this issue of *CUAJ*, Touma et al have analyzed the inter-rater reliability of scoring eight OSCE stations over two years of QUEST.¹ Not surprisingly, there is a variance in the scores given by two examiners assessing the same candidate at the same time on the same station. That said, I'm still surprised to see how well the stations performed, given all the possible distracting variables inherent in this form of candidate assessment. One wonders if the

poorer performance of the stations in 2020 reflects inexperience with virtual OSCEs at the beginning of the pandemic. By 2021, examiners and candidates were likely much more facile with Zoom, and the metrics presented suggest so.

Regardless, this type of variation in examiner grading is not new, as noted by the authors; however, to my knowledge, it is the first attempt to measure such variability in Canadian urology OSCE exams. Given this center's more than 25 years' experience administering QUEST OSCEs, it is likely their findings are translatable to the RCPSC OSCE. The results also speak to the need for, and philosophical intention of de-emphasizing examinations in the brave new world of competency-based medical education (CBME). To date, we have not seen this de-emphasis despite our specialty being 6–7 years into CBME.

The authors should be congratulated on their attempts to rigorously assess the validity of a unique and longstanding Canadian program.

COMPETING INTERESTS: The author does not report any competing personal or financial interests related to this work.

REFERENCE

1. Touma NJ, Peco CA, MacIntyre I. Inter-observer variance of examiner scoring in urology Objective Structured Clinical Examinations. *Can Urol Assoc J* 2024;18:116-9. <http://dx.doi.org/10.5489/cuaj.8571>

CORRESPONDENCE: Dr. Andrew MacNeily, Department of Urologic Sciences, University of British Columbia, Vancouver, BC Canada; amacneily@cw.bc.ca