PEER REVIEWED

Development of a Pilot Objective Structured Clinical Examination in Optometric Education

Patricia Hrynchak, OD, MScCH (HPTE), FAAO, Dipl AAO, Jenna Bright, BSc, MSc, OD, Sarah MacIver, BSc, OD, FAAO, Natalie Hutchings, BSc, PhD, MCOptom, Stanley Woo, OD, MS, MBA, FAAO, Dipl AAO

Abstract

Introduction: An Objective Structured Clinical Examination (OSCE) is a performance-based examination in a simulated environment intended to assess multiple clinical abilities or competencies. OSCEs have been shown to be a valid tool in health care education when appropriately developed. This paper reports on the use of an OSCE in optometric education at a North American School of Optometry and Vision Science. Methods: A pilot OSCE was developed and administered to a volunteer group of graduating Doctor of Optometry Students. There were 11 active stations and 3 rest stations. Six of the active stations involved standardized patients (SP). Skills were also demonstrated. Development, administration and post-OSCE considerations are discussed. There were 90 students eligible to take the OSCE. Of those, 54 volunteered and 53 consented to have their examination results reported. Of the 53 students, 46 (87%) passed the examination. Results: The OSCE was shown to be valid using Kane's theory on validity which looks at scoring, generalization, extrapolation and decision making. Conclusion: This pilot project demonstrated that an OSCE is a feasible assessment method to use in optometric education to prepare students for external certification processes. This paper shows educators that a strong development process including blueprinting with appropriate case development and training to standardize the examiners and SPs can produce a valid examination with appropriate test interpretation.

Key Words: *OSCE, objective structured clinical examination, assessment, education, optometry*

Introduction

In health care education and curriculum development, it is essential to have strategic alignment between learning objectives, teaching methods and assessment methods.[1] When the learning objectives involve developing a set of clinical competencies, the assessment methods should align with those objectives. The practice of optometry requires cognitive, technical and interpersonal skills and, as such, requires a coordinated system of assessment methods that go beyond multiple-choice questions and a straightforward demonstration of technical skills.[2]

An objective structured clinical examination (OSCE) is an assessment format where students rotate from one station to the next and are expected to perform a series of clinical tasks.[3] OSCEs are widely used in high stakes assessment in health care[4,5] OSCEs have been described in medicine, pharmacy, physiotherapy, massage therapy, nursing, midwifery and dentistry.[6] They have been used in certification assessments in Canada and the United Kingdom.

The University of Waterloo, School of Optometry and Vision Science developed and piloted an OSCE

with the aim of adding it to the assessment system used to determine end-of-program student competency. The results of the assessment were designed to inform future academic decisions and to give the students exposure to an assessment method used by the Optometry Examining Board of Canada (OEBC) for entry-to-practice eligibility.[7]

This paper reports on the feasibility of using an OSCE as an assessment tool in an optometry program through this pilot examination. It provides information to aid in developing and replicating the process within an optometric education context with best practice recommendations so that validity can be achieved.

Methods

A Learning Innovations and Teaching grant was awarded by the University of Waterloo, Centre for Teaching Excellence to fund the development and administration of the OSCE.[8] Research ethics approval was obtained from the Office of Research Ethics at the University of Waterloo, which follows the Declaration of Helsinki (See **Table 1**).

Step	Components
Preparation and Planning	• Organizational structure: OSCE examination team • Examination positioning within the curriculum • Blueprinting and examination length • Developing a bank of OSCE examination stations • Choosing a scoring rubric and standard-setting • Developing a pool of trained examiners • Developing a pool of trained SPs
Running the OSCE	• Choosing an objective OSCE venue • Setting up the OSCE circuit and equipment • OSCE circuit and troubleshooting
Post-examination considerations	• Handling results • The examination boards and ratification • Publication of the results • Complaints and appeals • Post-hoc psychometrics • Evaluation

Table 1. A summary of the organization and administration of an OSCE.[3] Click to enlarge

Preparation and Planning

The OSCE team consisted of four faculty members with a specific interest in assessment. An additional member was added during the analysis phase. The OSCE was proposed to be positioned in the curriculum as a new assessment to supplement the clinical assessment method used in the program's final year. The examination content blueprint was developed to reflect the competencies required by a graduating optometrist (as developed by the OEBC).[9] The practice areas of the competency profile are communication, professionalism, patient-centerd care, assessment, diagnosis and planning, patient management, collaborative practice, scholarship and practice management. The domains of competence thought to be best assessed were communication, professionalism, patient-centerd care, assessment (including psychomotor skills), diagnosis and planning and patient management. The content areas were refractive care, binocular vision and ocular disease. A two-dimensional matrix was created with the generic competencies along one axis and the content areas along the other (See **Table 2**). The weighting of the competencies being assessed were modeled on the weightings developed for the OEBC's examination using the frequency of a competency area in combination with the importance of that area, as determined by optometrists across Canada.[9]



Table 2. Blueprint: The content is blueprinted to the competencies to be measured. The “stars” are an individual measurable indicator as described by the Optometric Examining Board of Canada competency profile. The frequency of each indicator tested is dependent on the frequency and importance of the ability as judged by practicing optometrists. The Rigid Gas Permeable case (number 1) and allergic conjunctivitis (number 3) includes the exact indicators for the case as a examples. Click to enlarge

The length of the examination needed to be long enough to ensure the validity of the examination but short enough to be feasible.[3] We elected to use 11 active stations of 10-minute duration.[3] When including the three rest stations, the testing time for the examination was slightly less than 2.5 hours.

In developing a bank of OSCE stations consideration was given to the competencies that could realistically be assessed using the OSCE. The station writing was done by the team members.

Six of the 11 stations developed involved the use of SPs. The choice of having a SP depended on the competencies being assessed in that station. Not all competencies assessed required a SP. A station writing template was created based on the OEBC’s framework.[10] The components developed were the Case Information, Instructions to the Students, Summary of Patient Examination Record and Assessment Forms (See **Appendix A** for a sample). In addition, a comprehensive list of equipment, supplies and props was developed for each station, including the furniture and electrical sources required. Two team members piloted each station and adjustments to the stations were made to account for timing, clarity and difficulty level.

The development of a scoring rubric is critical for standardization. Scoring rubrics can be either analytical or holistic.[3] An analytical checklist is a series of statements outlining the performance that is expected. In alignment with what is generally accepted as appropriate in the literature, the psychomotor competencies (skills) were assessed using a checklist and the other competencies using global rating scales (holistic).[11] (See **Appendix B** for samples)

The passing score for the examination was determined using standard-setting.[12] This examination used the borderline group method to determine the passing score. The examiner rated the student’s overall performance on a station using a global rating scale holistically that included a borderline option independently of the grade from the scoring sheets. In stations where student performance was rated as

borderline on overall performance the average of the assessment scales was calculated and used as the pass grade for the station. This is different from arbitrarily setting a passing grade of, for example, 60%. The total of all borderline pass scores became the passing score for the examination.[11,12] In our examination, the overall performance scale was inadequate/poor/borderline/clear pass/outstanding.

The examiners attended a 3-hour training session to help standardize how the OSCE would be implemented and evaluated. The training covered 1) the principles of the OSCE, 2) the role of the examiners, 3) an explanation of the standard-setting method, and 4) an assessment of video-recorded OSCE stations[13], where three of the team members had been video recorded acting as the student, the examiner and the SP. The student was portrayed as either performing well, moderately, or poorly for one of the stations. The examiners viewed the videos and scored the student's performances. The scores were then shared between the examiners and discussed to try to develop a common framework for consistency.

An established SP program was contracted to provide the SPs. The trainer worked with the team to develop a script and train the SPs to portray the character in the clinical scenario for each station that used them. Each SP was asked to memorize their script and trained for approximately 1 hour. If more than one SP was used for the same station, they were trained together to ensure consistent performance. The SPs were experienced and compensated for their expertise.

Running the OSCE

We ran the OSCE in an optics laboratory hallway in adjacent rooms as clinic rooms were not available. We ensured that the rooms, additional classrooms and laboratories were booked exclusively for the OSCE for the 2 days. The examination was held between terms to decrease disruption. Rooms not used as stations were used for a briefing room for the students, space for examiners and SPs to report and administrative space to collect all scoring materials. A refreshment space was accessible, along with easy access to washrooms and water stations.[3]

Each station was set up in advance with the appropriate tables, chairs and equipment, ensuring sufficient space was available. In addition, any electrical or internet access needs were assured to be present. Each student was assigned a starting station number, which was clearly marked and easily seen on the door outside the station. We included the rest stations as starting stations for ease of coordination. The students moved in sequential order through the numbered stations until all stations were completed.

The examination instructions were posted outside the station on the wall in clear plastic protectors. In addition, the same instructions and other information were placed in clear plastic protectors and taped to the tables in the examination rooms to prevent writing on the materials or accidentally removing the materials after the end of the station.

Each examiner had station instructions and sufficient scoring sheets for the students. Recording forms were available and easily added to the binder with the scoring sheets after the student had finished recording. Each student had a sheet of identification stickers, which they handed to the examiner to affix to their recording sheet.

The required equipment was part of the documentation developed in the case writing stage.[3] Any new equipment was ordered and other equipment was determined to be in good working order well in advance of the examination. Back-up equipment was provided in case of malfunction or failure.

On the examination day, the students were given necessary instructions and expectations were reviewed. The SP and the examiners ran through their assigned station for a half hour before the examination started to ensure that they had similar expectations.

Movement of the students from one station to the next was managed by using a timed buzzer that was loud enough to be heard in the station rooms with the door closed. Before entering a station room, the student had 2 minutes to review the station instructions, at which time a buzzer sounded cueing the student to enter the room and begin. After 7 minutes, the buzzer sounded again alerting the student and examiner that there was 1 minute remaining. A double buzzer indicated the end of the station time. The students then moved to the next station. The examiners finished the assessment forms while the students were reading the instructions outside the room for 2 minutes. Hallway monitors were present to ensure that the students moved to the next station correctly, did not speak to one another and remained seated in a rest station.

Seclusion of the students could not be done, because we ran four circuits over 2 days. Ideally, we would have run two circuits simultaneously so that all students could be examined in 1 day. The students signed confidentiality agreements, agreeing not to disclose the content of the stations with other students. Security measures including barring cell phones and recording devices were in place.

Adverse event reporting forms were available for students to report if something happened that affected their performance. The forms included the name of the student, the station number, the date and time of the incident and a section to write what transpired. Instructions on how to submit it were also included. No forms were submitted by students.

Post-examination considerations

At the end of the examination all assessment sheets were collected and reviewed to ensure everything was complete before the examiners left. The data were entered manually into a spreadsheet. It was then coded for confidentiality, removing the student's name and replacing it with a numerical identifier in compliance with ethics approval.

If the examination had been part of the actual assessment system rather than a pilot, the team would have met to discuss the results and determine if they were acceptable. For example, was the pass rate adequate?

Of the 90 students who were eligible to take the examination, 54 volunteered and 53 consented to having their results reported. Of the 53 students, 46 passed the examination. The students did not receive results for this pilot as they needed to be analyzed first.

While not required for this pilot, a complaint process is necessary for an OSCE, to address any examination concerns or rescoring requests. If legitimate concerns are found regarding a specific station or if equipment malfunctioned, subsequent iterations of the examination can be informed. Also, this process could be used to reconsider scoring if necessary. We had developed an adverse event reporting form that could have been used if there was an incident.

In the post-hoc psychometric analysis of the OSCE results the reliability of the scores generated was determined. We used the Cronbach ?, which is a measure of internal consistency where the better students do well across all the items in the examination.[14] The Cronbach ? was 0.4 between stations. The interpretation is <0.50 unreliable, 0.50-0.80 borderline reliability, >0.80 reliable. However, all stations resulted in a lower ? when they were dropped from the analysis. Exploratory factor analysis identified three stations with negative factor loadings. For these stations, the individual student ratings were not well distributed across the global rating scale and had the poorest correlation with the overall OSCE score. A generalizability theory analysis could be added to the psychometrics to help define and attribute weight to various sources of error in the examination.[15]

The students were invited to give feedback on the examination. The student satisfaction survey results have been analyzed and published.[16] The overall satisfaction level was very high. Students were very

positive about the interactions with the SPs and the organization of the examination. They felt the examination used realistic clinical scenarios. There was mixed reaction to the use of simulators for skills assessment and the length of time available in each station to perform those tasks. Students were uncertain of how to prepare for the examination. The specific comments on the stations indicated that more examiner training is needed, especially with respect to the specific expectations in each station.

Results

Any test format, including the OSCE, does not have an inherent validity.[17] The understanding of validity has changed from considering separate types of validity to a single concept of construct validity.[18] In Kane's view, validity is a structured argument, in support of the interpretation of the score.[11] He has four components to this argument: scoring, generalization, extrapolation and decision making. The validity of this examination was evaluated using these components.

Scoring assures that there is evidence that the assessment data collected on each student (e.g., check sheets and global rating scales) have been scored accurately and collected appropriately. Also, the conditions of the examination should be standardized.[11] In this examination, the SPs were trained appropriately and monitored to ensure interpatient and intra-patient consistency. The examiners were trained appropriately in scoring methods and the assessment data were reviewed for missing information once the examiner submitted the forms. There was a threat to validity in that access to test content information was only controlled by having the students sign a confidentiality agreement. The score of a person who cheats on the examination is not representative of ability. Improvement in this aspect with sequestration of students who have completed the assessment should be considered.

Generalization focuses on the relationship between the observed scores and the true scores. The true score is the score that the student would receive if completing an unlimited number of assessments of the same type. It is important to remember that any test is only a sample of the content domain.[19] Therefore, the test items should be representative of the domain and consider the likelihood of obtaining similar scores if new items are used.[11,19] This examination achieved adequate sampling using an appropriate blueprint. The generalization concept also includes reliability or reproducibility of the numeric scores. As reported earlier, we used a Cronbach ? and achieved a reliability of 0.4. Reliability can be improved with more items, more stations, more raters or more occasions.[18] Indeed, if all students had taken the examination, the reliability may also have been higher. All of these aspects have to be balanced against feasibility constraints such as the size of the student pool, time, space and funding. Global rating scales might perform better than checklists.

Extrapolation inference looks at a sample of observations and generalizes it to the test-world universe, but how these scores represent real-world performance is paramount.[19] Cases should be developed that authentically represent the problem.[18] They should undergo review and piloting to ensure appropriateness to the students.[18] This examination achieved these goals and the student survey supported the view that the stations were authentic and would commonly be encountered in practice. Future work will look at how the OSCE scores correlate with other measures of performance throughout the program.

Inference speaks to interpretation of the evidence in making a decision.[18] This aspect of validity is achieved by appropriately determining the cut score or passing score for the examination.[18] In this test, the cut score was determined using the borderline groups method that is appropriate for this type of assessment.

Common challenges in utilizing an OSCE as an assessment format include the cost and the administrative and development time needed.[5,20] The cost includes direct costs such as paying SPs, buying equipment, supplies and catering.[4] Indirect costs are administrative and faculty time for

development and implementation.[4] In addition, administrative and faculty time may simply not be available. Another significant challenge is extremes in examiner judgement that result in a reduction in reliability of the examination.[19] Also, students find that the OSCE produces high levels of stress and anxiety.[22,23,24] Students may find that they are uncertain how to prepare for the examination especially when first exposed to the assessment format. A study by Müller et al.[25] showed that students should focus on skills labs and collaborative practice more than knowledge sources such as lectures and textbooks when preparing for an OSCE.

Conclusions

This project demonstrated that the OSCE is a feasible assessment method to use in optometric education. A strong development process can produce a valid examination. The currently preferred concept of validity as described by Kane explains that the collection of evidence is in support of the interpretations of test score.[11] There must be a structured and coherent argument that leads from the test administration to the interpretation.[11] “That structured argument is only as strong as its weakest component.”[11]

Our OSCE has carefully followed test development and administration protocol to meet this threshold for decision making. However, the examination results will allow for further refinement of the examination elements. It is recommended that attention be paid to developing faculty awareness and support, developing a pool of well-trained examiners, utilizing the abilities of the SPs well in advance of the administration of the examination and assuring that each station has sufficient time to complete the tasks.

References

1. Biggs J. Enhancing teaching through constructive alignment. Higher Education*:* 1996;32(3):347-64. doi: 10.1007/s10726-016-9285-4
2. Norcini J, Anderson MB, Bollela V, Burch V, Costa MJ, Duvivier R, et al. 2018 Consensus framework for good assessment. Med Teach*:* 2018; 40(11)1102-11. doi: 10.1080/0142159X.2018.1500016
3. Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The objective structured clinical examination (objective structured clinical examination): AMEE Guide No. 81. Part II: organisation & administration. Med Teach: 2013;35(9):e1447-63. doi: 10.3109/0142159X.2013.818634
4. Patrício MF, Julião M, Fareleira F, Carneiro AV. Is the objective structured clinical examination a feasible tool to assess competencies in undergraduate medical education? Med Teach: 2013;35(6):503-14. doi: 10.3109/0142159X.2013.774330
5. Casey PM, Goepfert AR, Espey EL, Hammoud MM, Kaczmarczyk JM, Katz NT et al. Association of professors of gynecology and obstetrics undergraduate medical education committee. To the Point: reviews in medical education–the objective structured clinical examination. Am J Obstet Gynecol: 2009;200(1):25-34. doi: 10.1016/j.ajog.2010.04.035
6. Hodges BD. The objective structured clinical examination: three decades of development. J Vet Med Educ. 2006 Winter;33(4):571-7. doi: 10.3138/jvme.33.4.571
7. Optometric Examining Board of Canada [Internet]; Taking the exam. [cited 2019, May 21] Available from http://www.oebc.ca.
8. University of Waterloo Centre for Teaching Excellence [Internet]: LITE Grant Projects. [cited 2019, June 11] Available from https://uwaterloo.ca/centre-for-teaching-excellence.
9. Cane D, Penny M, Marini A, Hynes T. Updating the competency profile and examination blueprint for entry-level optometry in Canada. CJO*:* 2018;(2)80:25-34. doi: https://doi.org/10.15353/cjo.80.267

10. Optometry Examining Board of Canada [Internet]: Case template. [cited 2019, May 21] Available from http://www.oebc.ca/written-osce/preparing-for-the-exam.
11. Holmboe E, During S, Hawkins R. Practical Guide to the Evaluation of Clinical Competence. 2nd Ed. Philadelphia, PA: Elsevier; 2018.
12. Newble D. Techniques for measuring clinical competence: objective structured clinical examinations. Med Educ: 2004;38(2):199-203. doi: 10.1111/j.1365-2923.2004.01755.x
13. Stanwick T, Forrest K, O'Brien B. Understanding Medical Education: Evidence, Theory, and Practice. 3rd Ed. Hoboken, NJ: Wiley; 2019.
14. Pell G, Fuller R, Homer M, Roberts T. International Association for Medical Education. How to measure the quality of the objective structured clinical examination: A Review of Metrics – AMEE Guide no. 49. Med Teach: 2010;32(10):802-11. doi: 10.3109/0142159X.2010.507716
15. Blood A, Park YS, Lukas R, Brorson B. Neurology objective structured clinical examination reliability using generalizability theory. J Neurol*:* 2015;85(18):1623–29. doi: 10.1212/WNL.0000000000002053
16. Hrynchak P, Bright J, MacIver S, Woo S. Student satisfaction with an objective structured clinical examination in optometry. J Optom Educ*:* 2021:46(2):27-32.
17. Hodges B. Validity and the OSCE. Med Teach*:* 2003;25(3):250-4. doi: 10.1080/01421590310001002836
18. Daniels VJ, Pugh D. Twelve tips for developing an OSCE that measures what you want. Med Teach: 2018;40(12):1208-13. doi: 10.1080/0142159X.2017.1390214
19. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: A practical guide to Kane's framework. Med Ed*:* 2015;49(6):560-75. doi: 10.1111/medu.12678
20. Kreptul D, Thomas RE. Family medicine resident OSCEs: A systematic review. Educ Prim Care*:* 2016;27(6):471-7. doi: 10.1080/14739879.2016.1205835
21. Fuller R, Homer M, Pell G, Hallam J. Managing extremes of assessor judgment within the OSCE. Med Teach*:* 2017;39(1):58-66. doi: 10.1080/0142159X.2016.1230189
22. Young I, Montgomery K, Kearns P, Hayward S, Mellanby E. The benefits of a peer-assisted mock OSCE. Clin Teach*:* 2014;11(3):214-8. doi: 10.1111/tct.12112
23. Brown J. Preparation for objective structured clinical examination: A student perspective. J Perio Pract: 2019;29(6):179-84. doi: 10.1177/1750458918808985
24. Bevan J, Russell B, Marshall B. A new approach to objective structured clinical examination preparation – ProOSCE. BMC Med Educ: 2019;19:126. doi: 10.1186/s12909-019-1571-5
25. Müller S, Koch I, Settmacher U, Dahmen U. How the Introduction of objective structured clinical examinations has affected the time students spend studying: Results of a nationwide study. BMC Med Educ: 2019;19(1):146. doi: 10.1186/s12909-019-1570-6

Appendices



Appendix A: Case description for case #1. Click to enlarge



Appendix B: Examiner scoring sheets as developed for station #1. Section A is an analytical checklist for a psychomotor skill. Sections B and C are holistic assessments (global rating scales) with the answers provided for the examiners to aid in grading. Click to enlarge

Dr. Patricia Hrynchak is a clinical professor emeritus at the University of Waterloo, School of Optometry and Vision Science. She has a master's degree in health practitioner teacher education and a Diplomate in Optometric Education from the American Academy of Optometry.

Dr. Jenna Bright is currently in independent private practice. At the time of this study she was the Director of the International Optometric Bridging Program, University of Waterloo, School of Optometry and Vision Science. Prior to completing her Doctor of Optometry, Dr. Bright completed a master's in science with a focus on patient centred communication in an optometric setting. Dr. Bright has a special interest in optometric education and cultural competency.

Dr. Sarah MacIver is a clinical associate professor at the University of Waterloo, School of Optometry and Vision Science with a special interest in the treatment and management of chronic ocular disease, specifically glaucoma and dry eye. She has been invited to speak at various education events including inter-professional conferences and public education events to disseminate knowledge about eye health and vision care. Areas of research include glaucoma, dry eye disease, inter-professional collaboration with primary healthcare and optometric education.

Dr. Natalie Hutchings is the Associate Director for Academics and Student Affairs at the University of Waterloo, School of Optometry and Vision Science. She has a special interest in optometric education. She works in the areas of ocular imaging, scaling and clinical assessment and spectacle lens adaptation.

Dr. Stanley Woo is the Director of the University of Waterloo, School of Optometry & Vision Science. He is a Diplomate in low vision with the American Academy of Optometry and Diplomate with the American Board of Optometry. His research interests include vision rehabilitation, ophthalmic imaging, public health policy and systems for optimizing patient care outcomes.