



The efficacy of peer assessment in objective structured clinical examinations for formative feedback: a preliminary study

Kyong-Jee Kim¹ and Giwoon Kim²

¹Department of Medical Education, Dongguk University School of Medicine, Goyang, and ²Department of Emergency Medicine, Soonchunhyang University Bucheon Hospital, Soonchunhyang University College of Medicine, Bucheon, Korea

Purpose: We sought to determine the impact of medical students' prior experience of assessing peers in the objective structured clinical examination (OSCE) on their clinical performance.

Methods: Forty-two year 4 medical students participated in an OSCE comprised of three 10-minute stations (syncope, hemoptysis, and back pain). Each student took part in two iterations of the three-station OSCE as either the examiner or examinee, and student performance was assessed using a checklist by a medical faculty member and a student simultaneously. Students were randomly assigned to two groups and their OSCE scores were compared. Students in the control group were tested at a station first and then participated at the same station as a peer examiner, and those in the intervention group participated as a peer examiner first and then as an examinee. Moreover, student OSCE scores rated by peer examiners were compared with those awarded by faculty to evaluate the accuracy of peer assessment. Following the test, students completed surveys on their perceptions of the usefulness of this formative OSCE.

Results: Student overall OSCE scores did not differ between groups. Students in the study group performed better at the hemoptysis station ($p < 0.001$), but poorer at the syncope station ($p < 0.01$). Student performances at the back-pain station were similar in these two groups ($p = 0.48$). OSCE scores rated by faculty and peer examiners were moderately negatively associated at the hemoptysis station ($p < 0.05$), but no such association was observed at the other two stations. This trend was similar in peer examiners who were high-achievers and low-achievers in OSCEs. Students showed positive perceptions of their experience with this OSCE.

Conclusion: Student experience as peer assessor offers a feasible means of providing them greater access to OSCEs without consuming more resources, although its impact on enhancing performance in the OSCE is likely to differ across stations.

Key Words: Medical students, Clinical competence, Educational measurement

Introduction

Objective structured clinical examinations (OSCEs) were devised to address a growing need in basic medical education to improve medical student competencies in

real-world situations [1]. The OSCE method is now well-recognized globally as a tool that substantively meets this requirement [2]. Past studies have shown that OSCEs provide a valid and reliable assessment tool when they are designed and implemented appropriately [1–3], although the reliability may differ across stations and

Received: December 2, 2019 • Revised: January 9, 2020 • Accepted: February 4, 2020
Corresponding Author: Giwoon Kim (<https://orcid.org/0000-0003-2720-7442>)
Department of Emergency Medicine, Soonchunhyang University Bucheon Hospital, Soonchunhyang University College of Medicine, 170 Jomaru-ro, Wonmi-gu, Bucheon 14584, Korea
Tel: +82,32,621,5300 Fax: +82,32,621,5016 email: kimgw@schmc.ac.kr

Korean J Med Educ 2020 Mar; 32(1): 59–65.
<https://doi.org/10.3946/kjme.2020.153>
eISSN: 2005-7288

© The Korean Society of Medical Education. All rights reserved.
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

depends on the number of stations and examiners [3].

Originally designed to enable experts to assess student competence, OSCE is also often employed to enable peer assessment and feedback [4]. OSCEs can be used for both formative and summative purposes [1], and several researches have been done on the effectiveness of OSCEs for formative feedback. Bernard et al. [5] found that medical students' review of formative OSCE scores were associated with their performance in subsequent summative OSCEs. Also, Daniels et al. [6] found from a study of internal medicine residents that having them immediately review their score sheets in OSCEs had a positive impact on their learning by giving timely feedback. Still, a study of year 2 UK medical students by Chisnall et al. [7] indicates that formative OSCEs are associated with improved performance in subsequent summative OSCEs only for identical stations.

Peer assessment is often employed in formative OSCEs. Peer assessment in OSCEs has the potential to promote learning by encouraging students to understand expectations and strategies for the test [4,8]. Peer assessment in OSCEs give students an opportunity to be exposed to the scoring rubrics, and research indicates that the disclosure of scoring rubrics in OSCEs enhance student test scores, particularly in history taking and physical examination scores [9]. A scoping review by Khan et al. [4] illustrated the benefits and challenges of peer assessment in OSCEs. Specifically, they found that peer assessment promotes learning, but the scores rated by students tended to be unreliable as compared with faculty scores, which suggested that students need to be trained on how to perform peer assessments in OSCEs. Moreover, research indicates that the accuracy of self-assessment of clinical performance is affected by such factors as gender, task familiarity, and years in medical study [10–12]. On the other hand, some studies have shown that medical students can act as effective peer exam-

iners [8,13–17] and that student-led OSCEs constitute a sustainable cost-effective approach [15].

Despite the effectiveness of peer assessment for formative OSCEs, students frequently have no opportunity to practice OSCEs other than in the high-stakes examination itself [8] and empirical evidence is lacking on the impact of student experience as peer examiners on their performance in OSCEs. We considered that student experience of OSCE as a peer examiner would benefit their OSCE performances by giving them opportunities to observe other students, therefore to learn from others and also enhance their understanding of scoring rubrics in OSCEs, which would help reflect on their performance and come up with better test strategies.

In this study, we explored the efficacy of a revised format of OSCEs, whereby students were exposed to each station twice, that is, once as an examinee and once as a peer examiner, by investigating the impact of this approach on student performance. In addition, we investigated the association between student OSCE scores assessed by standardized patients (SPs) and peer assessors to examine the accuracy of peer assessment. Summarizing, our study hypotheses were: (1) students who undergo OSCEs after experiencing the role of examiner perform better at that station than those without such experience; (2) students who perform better in clinical performance tests better assess the clinical performances of other students.

Methods

1. Study setting and procedure

The study participants were all of year 4 medical students (n=42) at a mid-sized private medical school in South Korea. Students participated in an OSCE com-

prised of three stations (syncope, hemoptysis, and back pain). Testing followed a general OSCE format, that is, each student was asked to perform in a simulated clinical situation by interacting with a SP by interviewing the patient, history-taking, conducting physical examinations, and arriving at a possible diagnosis and treatment plan. The time allocated at each station was 10 minutes. During this test, medical faculty members specializing in the field related to specific OSCE stations played the role of a SP. In this test, a peer examiner was also present in the station for assessment in addition to the SP. Student performance at each OSCE station was assessed by a SP and a peer examiner simultaneously. Students were assessed for patient interview and physical examination using a checklist format. The maximum possible score at each station was 100 points, which was an average of patient interview and physical exam scores. Both SPs and peer examiners used the same scoresheet.

For this experimental study, students were randomly divided into two groups. Students in the control group ($n=21$) were first tested at a station and then participated at the same station as a peer examiner, whereas students in the study group ($n=21$) first underwent the station as a peer examiner and then were tested at the same station. Only SP scores were counted towards student's final mark and the scores awarded by peer assessors did not contribute to the final assessment mark as the peer assessment was offered for formative purposes. Student OSCE scores were disclosed after a few days of the test. Students were surveyed regarding their perceptions of this formative OSCE format using a questionnaire comprised of three Likert-type questions and one open-ended question at the wrap-up session at the completion of the test. Students responded to the Likert-type questions using a 10-point scale, where 1 represented "strongly disagree" and 10 "strongly agree." The open-ended question solicited students' overall ex-

periences of the modified OSCE.

2. Data analysis

OSCE scores of students in the control and study groups were compared. The independent t-test was used to compare student OSCE score between groups. Pearson's r coefficients were used to establish relationships between student test scores as assessed by faculty and peer examiners to analyze reliability of peer assessment. To compare the accuracy of peer assessment across students who were high performers and low performers, the 42 students were dichotomized by overall OSCE scores into high- and low-achiever groups. The assessment scores allocated to peers by members of these two groups were compared with scores allocated by faculty SPs. Descriptive statistics was used to analyze questionnaire data. The analysis was performed using IBM-SPSS ver. 20.0 (IBM Corp., Armonk, USA) and statistical significance was accepted for p -values <0.05 .

3. Ethical considerations

Institutional review board (IRB) approval was not requested for this study, because it was part of the annual survey of students that pertain to their learning outcomes, which fell under the general exemption from our IRB for educational outcomes data. Participation was voluntary and consent was implied with the return of the survey as responses were collected anonymously.

Results

1. Student performance at objective structured clinical examination stations

Table 1 summarizes student performances at the three

OSCE stations for group A (tested first) and group B (acted as a peer examiner first). Overall OSCE scores were similar in these two groups ($t=0.562$, $p=0.575$). Group A performed better at the syncope station ($t=3.02$, $p<0.01$) but poorer at the hemoptysis station ($t=0.66$, $p<0.001$). Group performances at the back-pain station were similar ($t=0.71$, $p=0.48$).

2. Associations between standardized patient and peer examiner performance scores

Table 2 illustrates associations between student OSCE scores as allocated by SPs and peer examiners. OSCE scores awarded by SPs and peer examiners were moderately negatively associated at the hemoptysis station

($r=-0.33$, $p<0.05$), but no association was observed at the other two stations.

Table 3 summarizes associations between student OSCE scores as allocated by SPs and by peer examiners in the high- and low-achieving groups. Student OSCE scores awarded by high achievers were not associated with scores awarded by SPs at any of the three stations. Student scores of history-taking skills in the hemoptysis station awarded by low achievers were moderately and negatively associated with scores awarded by SPs ($r=-0.44$, $p<0.05$). Still, scores awarded by low achievers were not associated with SP awarded scores at the other two stations.

Table 1. Student Performances at Objective Structured Clinical Examination Stations

Stations	History-taking scores			Physical exams scores			Total scores		
	Control group	Study group	p-value	Control group	Study group	p-value	Control group	Study group	p-value
Syncope	77.3	61.2	0.004	75.6	66.1	0.190	74.0	66.7	0.004
Hemoptysis	84.8	78.3	<0.001	45.2	72.0	<0.001	60.0	76.3	<0.001
Back pain	83.4	84.2	0.428	80.1	83.6	0.672	83.9	81.8	0.481
Total	78.1	75.8	0.303	70.6	70.4	0.942	72.6	77.9	0.575

Students in the control group were tested at a station first and then participated at the same station as a peer examiner, whereas students in the study group were exposed to the station as a peer examiner first before they were tested in that station. The maximum possible score at each station was 100 points, which is an average of the sum of patient interview and physical exam scores.

Table 2. Associations between Student Objective Structured Clinical Examination Scores as Awarded by Standardized Patients and Peer Examiners

	Stations		
	Syncope	Hemoptysis	Back pain
Pearson's r	-0.042	-0.334*	-0.062

* $p<0.05$.

Table 3. Associations between Student Objective Structured Clinical Examination Scores as Awarded by Standardized Patients and High- and Low-Achieving Peer Examiners (Pearson's r)

	Stations								
	Syncope			Hemoptysis			Back pain		
	History taking	Physical exams	Total	History taking	Physical exams	Total	History taking	Physical exams	Total
High achievers	-0.221	-0.330	-0.328	-0.306	-0.251	-0.336	-0.471	-0.240	-0.408
Low achievers	0.213	-0.105	0.163	-0.442*	-0.219	-0.377	0.104	0.299	0.153

* $p<0.05$.

3. Student perceptions of the usefulness of formative objective structured clinical examination

Students generally agreed with the statement that experience of OSCE as a peer examiner would be of benefit in their own exams (mean=8.44, standard deviation [SD]=1.82), and disagreed with the statement that the presence of a peer examiner at an OSCE station made them feel uncomfortable (mean=2.44, SD=2.34). Several students reported they felt they received more feedback by acting as a peer assessor in this OSCE format. Some typical students' comments were as follows:

"I got far more feedback from this OSCE, because it allowed me to learn by watching other students and see how they do differently from me. I hope to have more opportunities to do OSCEs like this."

"OSCE scores had not been enough for me to explain what I did not do well. In this OSCE, I better understood my weaknesses by understanding what I was assessed in that station in more detail by seeing the assessment criteria in the scoresheet."

"It was helpful because I learned some expressions of other students when they talked with the patient that I felt were useful, but I had never used."

Discussion

Our study shows student exposure to OSCE stations as peer examiners may have a positive impact on their performance at some, but not all, stations, in terms of test scores. Student positive perceptions of this OSCE format indicate several benefits for learning. First, it is likely that student experience as peer examiner gave them an opportunity to improve their knowledge, skills

and attitudes in interacting with the patient by learning from how other students perform. Second, students were also likely to learn how to perform physical examinations more accurately from the experience of assessing their peers by being exposed to the scoring rubrics. These findings are consistent with the literature that supports the benefits of formative OSCEs [4,8].

Yet, only weak associations were found between assessment scores rated by faculty members and peer examiners regardless of whether peer examiners were high or low achievers. This finding suggests high-performing students do not necessarily better assess student performance in OSCEs than their low-performing peers. This is in line with the findings of previous studies that peer examiners are probably less reliable assessors than faculty members and that they should be trained to perform as effective peer assessors [4,16]. Therefore, our OSCE format may be adequate for formative assessments, but not for summative purposes. Nevertheless, student perceptions of this formative OSCE format were positive, which provides another evidence of its efficacy. Furthermore, there were a negative association between student OSCE scores as awarded by SPs and peer examiners in the hemoptysis station. It is speculated this finding is related to student task familiarity with the station. It can be assumed that the hemoptysis station was more difficult than others as the student OSCE scores were the lowest in that station. As previous studies indicates the accuracy of student self-assessment of clinical performance in OSCEs is related to task familiarity [12], student inaccuracy in assessment in the hemoptysis station may be related to task familiarity or difficulty of that station.

Although our study shows limited effectiveness of peer assessment in this OSCE format, it is still worthwhile to explore its feasibility and to enhance its effectiveness as it allows students to practice OSCEs more frequently

than in the conventional format. Offering medical students opportunities to practice OSCEs is limited by financial, faculty, and administrative constraints [15]. The described formative OSCE format allowed students to experience OSCE stations twice without increasing resource requirements, but its educational impact on student clinical performance probably differ across stations. Futures research is warranted on which factors in the OSCE format made differences in the outcomes across stations to enhance its educational impact. Future research is recommended for more in-depth understanding of what students learn from their experience of the OSCE as a peer examiner and whether and how it changes their behaviors that would result in an improvement in their clinical performance. Such studies will better inform us on how to make student experience of OSCEs as peer assessor a more effective educational opportunity.

We acknowledge that the present study was a preliminary study and thus has several limitations and warrant future studies. First, we examined student performance at three OSCE stations only, and the study was conducted using a small number of medical students in their final year. Future larger-scale study is warranted to enhance the generalizability of our findings. Second, we studied students who were in the final year in medical school. The impact of experience of peer-based assessment on student performance in OSCEs may be depended on time spent on a medical program, and thus, we suggest research be conducted on medical students during their earlier years.

ORCID:

Kyong-Jee Kim: <https://orcid.org/0000-0002-1936-5128>;
Giwoon Kim: <https://orcid.org/0000-0003-2720-7442>

Acknowledgements: The authors thank to professor

Bora Kang, department of emergency medical technician, Masan University for data gathering.

Funding: This work was supported by the Soonchunhyang University Research Fund.

Conflicts of interest: No potential conflict of interest relevant to this article was reported.

Author contributions: KJK and GK had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: all authors; acquisition, analysis, and interpretation of data: all authors; drafting of the manuscript: KJK; and critical revision of the manuscript for important intellectual content: all authors.

References

1. Khan KZ, Ramachandran S, Gaunt K, Pushkar P. The objective structured clinical examination (OSCE): AMEE guide no. 81. Part I: an historical and theoretical perspective. *Med Teach*. 2013;35(9):e1437-e1446.
2. Patrício MF, Julião M, Fareleira F, Carneiro AV. Is the OSCE a feasible tool to assess competencies in undergraduate medical education? *Med Teach*. 2013;35(6):503-514.
3. Brannick MT, Erol-Korkmaz HT, Prewett M. A systematic review of the reliability of objective structured clinical examination scores. *Med Educ*. 2011;45(12):1181-1189.
4. Khan R, Payne MW, Chahine S. Peer assessment in the objective structured clinical examination: a scoping review. *Med Teach*. 2017;39(7):745-756.
5. Bernard AW, Ceccolini G, Feinn R, et al. Medical students review of formative OSCE scores, checklists, and videos improves with student-faculty debriefing meetings. *Med Educ Online*. 2017;22(1):1324718.
6. Daniels VJ, Strand AC, Lai H, Hillier T. Impact of tablet-scoring and immediate score sheet review on validity and

- educational impact in an internal medicine residency objective structured clinical exam (OSCE). *Med Teach*. 2019;41(9):1039-1044.
7. Chisnall B, Vince T, Hall S, Tribe R. Evaluation of outcomes of a formative objective structured clinical examination for second-year UK medical students. *Int J Med Educ*. 2015;6:76-83.
 8. Young I, Montgomery K, Kearns P, Hayward S, Mellanby E. The benefits of a peer-assisted mock OSCE. *Clin Teach*. 2014;11(3):214-218.
 9. Chae SJ, Kim M, Chang KH. Can disclosure of scoring rubric for basic clinical skills improve objective structured clinical examination? *Korean J Med Educ*. 2016;28(2):179-183.
 10. Madrazo L, Lee CB, McConnell M, Khamisa K. Self-assessment differences between genders in a low-stakes objective structured clinical examination (OSCE). *BMC Res Notes*. 2018;11(1):393.
 11. Blanch-Hartigan D. Medical students' self-assessment of performance: results from three meta-analyses. *Patient Educ Couns*. 2011;84(1):3-9.
 12. Fitzgerald JT, White CB, Gruppen LD. A longitudinal study of self-assessment accuracy. *Med Educ*. 2003;37(7):645-649.
 13. Burgess A, Clark T, Chapman R, Mellis C. Senior medical students as peer examiners in an OSCE. *Med Teach*. 2013;35(1):58-62.
 14. Chenot JF, Simmenroth-Nayda A, Koch A, et al. Can student tutors act as examiners in an objective structured clinical examination? *Med Educ*. 2007;41(11):1032-1038.
 15. Lee CB, Madrazo L, Khan U, Thangarasa T, McConnell M, Khamisa K. A student-initiated objective structured clinical examination as a sustainable cost-effective learning experience. *Med Educ Online*. 2018;23(1):1440111.
 16. Basehore PM, Pomerantz SC, Gentile M. Reliability and benefits of medical student peers in rating complex clinical skills. *Med Teach*. 2014;36(5):409-414.
 17. Moineau G, Power B, Pion AM, Wood TJ, Humphrey-Murto S. Comparison of student examiner to faculty examiner scoring and feedback in an OSCE. *Med Educ*. 2011;45(2):183-191.