Original Article
_____

Evaluate specialist faculty examiner and student-examiner of graduates performances on a preliminary osce

¹Ziyad Afzal Kayani, ²Sarmud Lateef Awan, ³Irum Gilani, ⁴Shagufta Manzoor, ⁵Farzana Sabir, ⁶Zakir Naqvi

¹Department of Surgery AJK Medical College, Muzaffarabad,
²Department of Public Health & ³Community Medicine, Health Services Academy, Islamabad,
⁴Department of Anatomy AJK Medical College, Muzaffarabad, Pakistan

Correspondence

Ziyad Afzal Kayani, Department of Surgery AJK Medical College, Muzaffarabad

Abstract

Aim: Medical postgraduates have already been utilized as ratings in realistic organized medical tests to assist relieve load on faculty personnel. There has been rare research on their capacity to fill lists and worldwide rating scales, although there is lack of evidence on their opportunity to offer evaluation to younger colleagues. The goals of our current research remained about to evaluate proficient facility inspector and student-examiner assessments of postgraduates' (applicants') behavior on a summative OSCE; (ii) evaluate SE comments offered to applicants; and solicit input on acceptability across all respondents.

Methods: Year 3 medical postgraduates (n = 74 applicants) took part in a 9-SF OSCE. Year 3 postgraduates (n = 29) worked as SEs, while teaching doctors (n = 29) worked as FEs. SEs and FEs separately rated applicants for each location employing specifications and worldwide rating scales. After each interaction, the station effectives offered comments to the applicants. Our current research was conducted at AJK Medical College Muzaffarabad from May 2020 to April 2021. The FEs graded SEs on response supplied using just the standardized rating scale (1 = strongly disagree, 5 = strongly agree) for numerous criteria, including how the feedback was equal, specific, correct, suitable, professional, and equivalent to responses given by the FE. All postgraduates completed out surveys on their opinions and acceptance. **Results:** On the requirements gathered, there would be a significant correlation among raters for every station, ranging from 0.58 to 0.87. Comparisons on worldwide rating varied from 0.24 to 0.79 for each station. Station effective response remained well respected by professor examiners, through mean ratings ranging from 5.03 to 5.45 crossways altogether areas. Candidates and examiners together articulated the strong level of getting.

Conclusion: In relation to the ability to fulfill checklists and offer comments, student-examiners seems to be the feasible substitute to FEs in the formative OSCE.

Keywords: Faculty Examiner, Student-Examiner, Evaluations of Graduates Performances, OSCE

INTRODUCTION:

The main structured medical exam is commonly utilized to determine medical post graduates' competence for together influential and summative objectives. Though OSCE scores might remain an accurate and effective evaluation of an examinee's medical competence, they place a significant burden on doctor efforts. Doctors are frequently used as examiners since they remain seen as "expert" judges, and several medical schools utilize doctors to offer response to applicants throughout the test procedure [1]. Due to the difficulty of obtaining doctor examiners, alternate ratings agencies, including such patients aged, have been used. Some other possible replacement to physicians, and one that we are interested in, is employment of studentexaminers. In medical school, medical post graduates have now been examined as assessors of other medical post graduates in a variety of circumstances [2]. Medical post graduates have now been demonstrated to rate lectures in addition written papers on past of medicine in same way as faculty members do. Arnold et al. noticeable in medicinal places that noble estimations however in an internal medication rebellion accessible fortunate reliable, unbiassed, and actual. Chenoa et al. inspected student-teachers' volume to resolve as OSCE assessors on past places in an OSCE context [3]. Graduate tutors outperformed teaching doctors on general (alterations of 0.03–0.21 on the 5-point Likert scale). Interrater confidence ranged from 0.42 to 0.65 for questionnaire valuation and worldwide ratings at stations, with a total interrater consensus of 0.67 on the final grade. Examinees (65%) expected similar grades from student and ability inspectors, and 86 percent expected SEs to be as objective as professor investigators [4]. Ogden et al. initiate that dental undergraduate remained consistent to dentists as examiners on mouth review station while in an OSCE once assessment to dental college workers. Despite the likely cost savings of utilizing SEs, it would be interesting to investigate their usage in a medical student OSCE assessing both history and clinical examination abilities, as well as their capacity to offer feedback. Furthermore, a more thorough examination of SE comments by careful observation through academic staff could considerably contribute to existing research. The goal of this research remained to assess specialized FE and SE assessments of scholars' performance on a formative OSCE. Even if researchers remain absorbed in comparison to summative tests among rater types, the primary goal of our current research is to evaluate prenatal assessment results [5].

METHODOLOGY:

Year 2 medical post graduates were given a formative nine-station OSCE. All medical post graduates in Years 2 and 4 were asked to take part. The first 69 Year 2 pupils who willingly participated as test - takers (applicants) were chosen from a pool of 154. Post graduates have never even taken an OSCE before and were given a 23-minute candidate education before to the exam. SEs were chosen from among the first 28 Year 4 student volunteers. They had already taken three OSCEs as aspirants throughout medical school. Maximum academic examiners were competent physicians who've been interested in education and had previously examined in at minimum one mandatory OSCE or OSPE in basic sciences. The OSCE consisted of eight stations that tested medical abilities such as history-taking, conversation, counseling, and physical examination. At separately location, a SE and FE saw the interaction concurrently and individually performed a checklist and a worldwide rating scale. The worldwide grading scale employed a 6-point scale with 1 being inadequate and 6 being exceptional, with 3 being borderline undesirable and 4 being borderline good. And though the global rating is intended to represent entire excellence of instructors' performance, it is only used to determine a cut score, then it is not taken into account when determining a station's score. The modified borderline group approach was employed as the standard-setting procedure. Our current research was conducted at AJK Medical College Muzaffarabad from May 2020 to April 2021. The mean

inventory score for all applicants designated as borderline is used to calculate the cut score to every station. In parallel to rating the applicants, FEs used a standardized evaluation form to rate SEs on verbal comments they provided to another applicant. In the prior OSCE, the form had just been tested. Post graduates completed the questionnaire in the allotted time and did not feel it had a negative influence on their grades (data not included). The 5-point rating system was being used to assess many elements of the feedback, including whether is balanced, specific, correct, and suitable. The FE had also been requested to score comments on a scale of one to ten, and to indicate if he or she might well have offered similar criticism and intervened, if necessary, as well as grade SE professionalism. Except if the FE felt it was necessary to interfere, the applicant received no more input from the FE. Three similar tracks were operated consecutively for two semesters to suit all pupils. Candidates cycled between stations while the SEs stayed immobile throughout the assessment. To prevent accidental the same SE and FE duo for the duration of the exam, the FEs switched tracks halfway through the exam. Participants' opinions on whether or not the procedure was satisfactory, as well as their impressions of it, were elicited using questionnaires.

RESULTS:

Table 1 shows average scores, margins of error, and relations among SEs also FEs for the diagnostic and global rating scales. SEs had an average overall checklist score of 7.93, whereas FEs had an average total score of 7.79 ($F_{1,67} = 8.23$, $p = 0.004$, $d = 0.38$). By way of seen in table, normal checklist score provided by SEs on every location expected to remain developed than equivalent checklist score offered through FEs. The mean global ratings of the SE and FE groups do not vary (means = 5.82 and 5.84, accordingly; F 1, d = 0.05). The associations among of checklist scores and the worldwide ratings offered by every rater type are also shown in Table 1. For average overall checklist score ($r = 0.83$), the correlation across rater types was strong ($r = 0.83$), and for individual station scores, it varied from reasonable to high ($r = 0.57$ to $r = 0.87$). For the mean worldwide rating ($r = 0.69$), the correlation across rater types significantly lower ($r = 0.69$), and for specific station ratings, it varied from 0.24 to 0.79. To evaluate the reliability of the two methods, a usability study was performed. The recurrent-measures descriptive analysis was used for list, to stations (1–9) being preserved as the repeated actions influence and rater type being cuddled inside every station. Similar research remained carried out for universal rating system. The variability factors for specification in addition global rating scale are summarized in Table 2. Though SP remained disordered by station, difference credited to SPs might not remain comprised inside moreover study. Aspirants contributed for almost nothing of range in scores for each measure, as seen in the table, suggesting that the cohort were presumably very homogenous.

Table 1:

Station No:	FE mean,	SE mean	Correlation for checklist	FE mean, global	SE mean, global	Correlation for global
1 HX	4.69 (0.90)	4.55 (0.93)	0.63	6.65 (1.02)	6.45 (1.07)	0.67
2 PE	4.95 (0.85)	4.89 (0.95)	0.58	6.91 (1.53)	6.99 (1.33)	0.87
3 HX	4.90 (0.82)	4.79 (0.69)	0.25	6.13 (0.98)	6.48 (0.93)	0.55
4 PE	5.09 (0.84)	5.02 (0.94)	0.79	7.03 (1.23)	7.08 (1.41)	0.83

5 PE	4.75 (0.87)	4.74 (0.88)	0.58	7.06 (1.30)	6.30 (1.39)	0.68
6 HX	4.48 (1.40)	4.92 (0.75)	0.23	7.26 (1.30)	7.59 (1.15)	0.81
General mean	4.82 (0.46)	4.81 (0.40)	0.69	6.78 (0.57)	6.91 (0.52)	0.84

Table 2:

Factor	Global rating			Checklist		
	Variance	%	SE	Variance	%	SE
Candidates	0.08	11	0.04	0.13	6	0.06
Stations (s)	0.01	1	0.01	0.46	24	0.24
Raters (r): s	0.02	2	0.01	0.08	4	0.04
cs	0.34	37	0.04	1.07	49	0.08
cr:s	0.44	49	0.03	0.41	19	0.02

DISCUSSION:

One reason for this consequence though once doctor-examiners exploits rating scale, they also might give an examinee advantage of hesitation meanwhile they understand what examinee remains attempting to do, before they might just penalize an examinee when somewhat was not done correctly [6]. The non-doctor examiner will be unable to comprehend what the examinee is doing due to a lack of medical expertise, in addition henceforth would be unable to credit or punish the examinee [7]. Through only the checklist, an examinee either does or does not do anything, accordingly the physician is just less likely to credit an activity. It is indeed likely this because when positively influenced are emphasized, worldwide scores become less unbiased [8]. This is undoubtedly one of Rothman and Cimiano's findings. The associations for two of the locations in specifically were lower in our analysis. Both stations took histories; station 3 featured the dysmenorrhea record, while station 6 required a backache history [9]. Station 1 remained the psychiatric history-taking station, nevertheless there was no evidence of a decreased rater relationship; hence, it is unclear to what extent good communication skills modify the connection among SEs and FEs on a rating scale [10].

CONCLUSION:

In conclusion, in a formative OSCE, SEs seemed to be a feasible option to FEs in based on its ability to execute criteria and receive suggestions.

REFERENCES:

1. Benbassat J, Baumal R. Uncertainties in the selection of applicants for medical school. Adv Health Sci Educ. 2017;12(4):509–21.
2. Patterson F, Ferguson E, Zibarras, L. Selection into medical education and training. In: Understanding Medical Education: Evidence, theory, and practice. Third edition. Edited by Swanwick T, Forrest K, O'Brien BC. Oxford: Wiley-Blackwell and The Association for the Study of Medical Education; 2019;375–88.

3. Kulatunga-Moruzi C, Norman GR. Validity of admissions measures in predicting performance outcomes: the contribution of cognitive and non-cognitive dimensions. Teach Learn Med. 2020;14(1):34–42.

4. Patterson F, Knight A, Dowell J, Nicholson S, Cousans F, Cleland J. How effective are selection methods in medical education? A systematic review. Med Educ. 2019;50(1):36–60.

5. Eva KW, Rosenfeld J, Reiter HI, Norman GR. An admissions OSCE: the multiple mini-interview. Med Educ. 2018;38(3):314–26.

6. Knorr M, Hissbach J. Multiple mini-interviews: same concept, different approaches. Med Educ. 2018;48(12):1157–75.

7. Trapmann S, Hell B, Weigand S, Schuler H. Die Validität von Schulnoten zur Vorhersage des Studienerfolgs – eine Metaanalyse [The validity of school grades for academic achievement – a meta-analysis]. Z Padagog Psychol. 2017;21(1):11–27.

8. Ferguson E, James D, Madeley L. Factors associated with success in medical school: systematic review of the literature. BMJ. 2020;324:952–7.

9. Albanese MA, Snow MH, Skochelak SE, Huggett KN, Farrell PM. Assessing personal qualities in medical school admissions. Acad Med. 2019;78(3):313–21.

10. Detailergebnisse des Zentralen Verfahrens (ZV) zum Wintersemester 2020/21 [Detailed results of the central selection procedure for the winter semester 2020/21]. https://www.hochschulstart.de/unterstuetzung/downloads#c2551. Accessed 31 Mar 2021.