Original Article

The Effect of Feedback on Objective Structured Clinical Examinations: A Quasi-experimental study

Mahdi Biglarkhani[1], Azam Meyari[1*], Hossein Emad Momtaz[2], Ali Reza Soltanian[3], Nasrollah Pezeshki[2], Mahnaz Yavangi[4], Gholamreza Gorbaniamjad[5]

[1] Persian Medicine Department, Medical School, Hamadan university of Medical Sciences, Hamadan, Iran.
[2] Pediatric Department, Medical School, Hamadan university of Medical Sciences, Hamadan, Iran
[3] Biostatistics Department, Public Health School, Hamadan university of Medical Sciences, Hamadan, Iran.
[4] Gynecology Department, Endometrium and Endometriosis Research Center, Medical School, Hamadan university of Medical Sciences, Hamadan, Iran.
[5] Orthopedics Department, Medical School, Hamadan university of Medical Sciences, Hamadan, Iran

Article Info

Article history:
Received 10 Apr 2020
Accepted 4 Feb 2021
Published 10 Mar 2021
Keywords:
Objective Structured Clinical Examination,
Clinical assessment,
Feedback,
Medical education

***Corresponding author:**
Azam Meyari, Persian Medicine Department, Medical School, Hamadan university of Medical Sciences, Hamadan, Iran.
Email: d_a_meyari@yahoo.com

Abstract

Background & Objective: Todays the Objective Structured Clinical Exams (OSCE) are used as new and valid tools for assess clinical skills of medical students and residents. Considering the importance and essentiality of these exams, this study was designed and conducted to assess effect using feedback for correcting the weaknesses and promoting the positive points of designing and performing this type of clinical exams.

Materials and Methods: This Quasi-experimental study was performed by using a valid and reliable researcher made questionnaire consist of principles of the OSCEs. At the first phase, the10 residency's OSCEs were observed and completed questionnaires by researchers and randomly divided two groups: control and witness. At the second phase the feedback of exams presented to faculty OSCE designers in control group. At the third phase the next OSCEs that designed by two groups were observed and evaluated. Data gathered entered in SPSS-15 and results compared by t-test and Paired t- test.

Results: After assimilation scores of two groups before intervention, there was significant statistical difference between before and after intervention in intervention group (pediatrics, gynecology, anesthesiology, internal medicine and pathology) (P<0.003) but there wasn't significant difference in witness group.

Conclusion: Using feedback confirm this technic by the experts in medical education can correct and promote quality of designing and performing of OSCEs.

Introduction

At the present, the common evaluations in the teaching-learning process are often the pivotal response, and even though clinical skills and practical work play the main role in medical education, the success of medical graduates largely depends on memorized materials and focuses on remembering simple information and doing low-level assignments. Meanwhile, high cognitive levels are essential in problem-solving and illness situations, which necessitates the use of novel assessment methods (1). In this regard, one of the new student performance assessment methods is known as the objective structured clinical examination (OSCE), which was first introduced by professor Harden (1975) in Scotland. However, this test is not similar to other tests that include oral, descriptive, or multiple-choice questions. In fact, the test has an organized framework consisting of several stations using a simulation or laboratory environment, the content of which is determined based on the objectives of the course. The test takers show their level of ability and skill during the specified time at each station while being assessed by trained supervisors at each station, who used predeveloped checklists based on the specific goals of each station.

These tests have played an important role in the evaluation of medical students' practical skills in the

past few decades (2). Most clinical professors believe that the mentioned test can better determine students' performance, compared to traditional tests (3-6). In a research by Manogue & Brown (1998), the OSCE was an effective feedback technique for students (7). In addition, the OSCE has a structured and completely defined scoring method, the details of which are determined based on principles and scientific formulas of medical education. All test designers must reach a consensus over various issues, including deciding about the content of each station, the checklist related to each station, scoring in each station, the minimum pass level of each station, and the total test quorum. During the test, students show their knowledge and skills in a simulated environment, and their level of preparedness for dealing with real-world situations is assessed during the process. Despite the effectiveness of OSCE in terms of the evaluation of medical students' clinical competency, various limitations such as experienced human resources, training people as standardized patients, resources, facilities and time-consuming processes have caused various executive problems, which necessitate accurate planning by professors (designers) and supervisors, the provision of a suitable place, proper measurement tools, and the necessary tools in each station for more success (8, 9).

In a critical review study, Turner et al. concluded that despite the advantages of the OSCE, attempts should be made to increase the number of these benefits and improve the reliability of the test. According to these scholars, a successful OSCE is the result of purposive, coordinated and committed planning to hold a large-scale test that can effectively evaluate individuals (10). In another study, the searchers mentioned the need for an experienced team to successfully hold an OSCE (11), affirming that a lack of access to these individuals, who can train standardized patients and develop clinical scenarios, was the main disadvantage of the test (12). Despite the scientific proofing of the effective role and importance of the OSCE, a low number of domestic studies have focused on the improvement of the quality of the test to get closer to scientific standards. Therefore, it seemed that performing the necessary interventions to improve the quality of design and implementation of these tests can be effective in increasing the validity and stability of their results. In addition, using feedback in its scientific sense (i.e., objective, conscious, and without judgment of) on the performance of individuals with the goal of improving skills presented by technical and special methods (13) can be effective in eliminating the shortcomings and problems of design professors and performers of the OSCE in residency courses in the next rounds. With this background in mind, the present study aimed to improve the quality of OSCE design and implementation in medical residency programs at the School of Medicine, Hamedan University of Medical Sciences using interventional feedback.

Materials and Methods

This quasi-experimental research was performed on 10 educational departments with residents selected by the census sampling method. First, we extracted the scientific and technical principles required for the design and implementation of OSCE by studying valid resources, followed by preparing a primary checklist in line with research goals. Afterwards, the validity of the checklist was confirmed by asking the opinions of experts in the field of medical education and teachers involved in these tests. In addition, the inter-rater reliability of the tool was approved by holding several coordination meetings and reaching consensus among the researchers. The researcher-made checklist encompassed two sections; the first part included information about the field of study, date and location of the test while the second section contained 24 questions about how to adhere to standard principles in the design and implementation of the OSCE. In this regard, five questions were related to the physical space of the test, whereas six, four, three, five, and one questions were about content of stations, a checklist of each station, specified time, scoring

method of the test, and the number of stations, respectively. The questions were answered based on a range of yes, to some extent, no. Since there was no standard for division, the scores were turned into a 100 scale with the agreement of the researchers. In this regard, scores were divided into above the third quartile (75%), between the second and third quartile (50-75%), and below the second quartile. If more than 75% of the stations of a test had the desired criterion, the answer was "yes" (with a score of 3) and if less than 50% of the stations had the desired criterion, the answer was "no" (with a score of 1). Moreover, if 75% -50% of the stations had the desired criteria, the answer was "to some extent" (with a score of 2). Therefore, the scores ranged from 24 to 72.

To perform the research, we first obtained the necessary licenses and made the required coordination with the school of medicine. We entered all 10 schools (four surgery and six internal departments) that held the OSCE at the residency level. The desired checklist was filled by the researchers after visiting and observing the OSCE held in each of the mentioned departments with former notification of the department and test developers. In order to adhere to ethical considerations, informed consent was obtained from the heads of the related departments. After that, the educational departments were randomly divided into five groups of control (surgery, dermatology, ear, nose and throat, cardiology, and psychology) and intervention (internal medicine, anesthesiology, obstetrics and gynecology, pediatrics, and pathology) departments, in a way that each category included two surgery and three internal departments. In the next step, appropriate and necessary feedback was provided to the test designers of the intervention group in separate meetings and coordination with the dean and department heads based on the observations and the contents of the relevant checklist, and the necessary educational consultations were designed and presented for them. The meetings were held one month before the next test and at the location of the Medical Education Development Center of the university. This feedback is based on the accepted principles in the science of medical education and in the form of objective, informed and non-judgmental critique of the performance of test designers and test-takers in the intervention group and by observing its conditions- i.e., direct, respectful, timely and with recommendations for improvement (13) made by experienced and proficient people in the field of medical education and the design and implementation of the OSCE test (preparation of test specification table, content of each station, preparation of checklists, allotted time, preparation of residents, optimal physical space, test management, and scoring method), as well as the necessary feedback on each station and related content. In the end, the necessary educational counseling sessions were held to remove defects depending on the limitations of each department, the spatial and temporal conditions, and the number of residents. On average, four hours of feedback provision were allocated to each department. In addition, an example of standard OSCE was exhibited in the form of separate slides. In the third stage and after the intervention, the researchers attended the next OSCE and collected data by filling the related checklists. Data analysis was performed in SPSS version 15 using descriptive and analytical statistics (e.g., paired t-test). The relevant principles were compared before and after the intervention in both intervention and control groups, followed by preparing the final report. It is worth noting that the present research was approved by the research council of the university and the ethics committee (letter No.: D/P/16/35/9/3237). In addition, the results were published anonymously and the researchers only referred to the educational departments.

Results

The mean scores achieved for each item in each group before and after the intervention are shown in Table 1. As observed, the scores of the intervention group

increased significantly after the intervention, while no significant difference was observed in the control group before and after the intervention (0.003 and 0.456, respectively).

Table 1: Total scores of groups before and after feedback

	group	Score before feedback	Score after feedback	P value
Intervention groups	Pediatric	53	65	0.003
	Gynecology and Obstetrics	44	70	
	Internal Medicine	50	68	
	Pathology	53	70	
	Anesthesiology	50	59	
Control groups	Dermatology	41	41	0.456
	Cardiology	41	40	
	Psychiatry	54	55	
	Surgery	51	52	
	ENT	59	45	

According to the results, the highest error rate in both groups before and after the intervention was related to adherence to points related to checklist preparation and scoring in each station, whereas the lowest score was related to the number of stations and allotted time of each station. There was no significant difference between the groups in terms of the scores of the items, except for the scoring method, which was higher in the control group, compared to the intervention group (Table 2).

Table2:The mean score of items in witness and intervention groups before feedback

	Intervention group Mean±SD	Witness group Mean±SD	P value
Aim & Content	12.4±1.5	12.4±2.5	1.00
Checklist	6.8±1.09	6.0±2.0	0.455
Environment	11±2	10.6±3.9	0.844
Time	8.2±1.3	8.4±0.54	0.760
Number of stations	2.4±0.89	2.4±0.89	1.000
Scoring	8.2±0.44	9.0±0.000	0.004
Sum	49±3.3	49.2±8.01	0.960

In addition, the two groups were homogenous regarding the scores obtained before the intervention, in a way that the t-test results demonstrated no significant difference in the mean scores of the two groups in this respect (P=0.96). After feedback, however, a significant difference was detected between the two control and intervention groups (P<0.001), and the items are presented in Table 3 separately. As observed, there was a significant difference between the two groups in all items, except for the allocated time and the number of stations.

Table3: The mean score of items in witness and intervention groups after feedback

	Intervention group Mean±SD	Witness group Mean±SD	P value
Aim & Content	16.6±0.89	11.4±1.5	0.000
Checklist	10.2±2.4	6.0±2.8	0.037
Environment	13.6±0.89	9.2±3.5	0.028
Time	9.0±0.000	8.6±0.054	0.141
Number of stations	3.0±0.000	2.4±0.89	0.172
Scoring	14.0±1.4	9.4±8.9	0.000
Sum	66.4±4.6	46.8±6.4	0.001

Discussion

The present study aimed to present feedback as one of the most important and emphasized strategies in medical education in its scientific sense, which is objective, conscious, and without judgment of the performance of the individual or individuals with the aim of improving skills in an appropriate and timely manner with an emphasis on enhancing the strengths and eliminating the weaknesses. Intervention through feedback increased adherence to the items of principles related to the design of a standard checklist and development of proper content for stations, as well as providing a suitable physical space for stations and using the scientific and standard formulation for scoring in the intervention group, showing a significant difference compared to before the intervention. There was no significant difference in the intervention group regarding the number of stations and appropriate test time. In this regard, our findings in the first four items are in line with foreign studies by Turner and Holmbole and domestic research of Noohi and Malekanrad (10-12, 14) regarding the necessity of using expert station designers, developing scenarios, employing skilled people to hold the test, and improving the quality of the stations' content. We also found no significant difference in the group regarding the last two items after the intervention, which might be due to the appropriateness of the stations and their durations in all educational groups before the intervention. In addition, there was no need for interventional feedback to designers in order to increase or decrease the number of stations and duration of each of them, standardize the process and make the necessary corrections. In a similar research, Hosainrezaee et al. improved the OSCE design for nursing students and increased its validity and reliability by correcting the education and test environment for the faculty, students, and those involved in the test (15).

Other studies employed different methods (e.g., focus group, Delphi method, and a pilot study) to improve the OSCE (16, 17). In another study by Garcia et al. (2017), the test results were analyzed in addition to holding the annual OSCE in pediatrics, which helped identify the weaknesses of students and shortcomings of the test and improve the process (18). In a review study, Adib et al. described the effect of the OSCE on nursing students' anxiety, satisfaction, and learning, concluding that students could be accurately analyzed by the accurate design of stations, proper preparation of students and teachers during education, providing the suitable environment and necessary equipment and workforce, and allocating sufficient time to the OSCE (19). Emadzadeh et al. (2017) performed a research at Mashhad University of Medical Sciences to evaluate the challenges of the OSCE from the perspective of the participants. According to the results, the subjects mainly complained about the disproportion of the time allotted to each station to the requested subject and

items. In addition, they marked that taking the same OSCE test during education could have a great impact on reducing participants' stress and their success in the pre-board OSCE (20). In a study by Shahzad et al. (2017), data were collected through interviews with dental students and graduates in Pakistan. Students mostly complained about the insufficient time of stations and lack of coverage of all educational goals (21). Contrary to the recent research, these studies recognized the lack of proportion of the time of each station to its content as a weakness of the OSCE. This could be due to the standard allocation of time to each station in the present study and evaluation of the issue from the perspective of students and test-takers in the mentioned studies.

In a study conducted by Humayun (2016), it was concluded that despite many studies and measures taken to improve OSCE tests, they need to be upgraded in terms of scoring and educating standard and simulated patients, as well as training evaluators at each station (22). In another qualitative study conducted by Obizoba through observation and semi-structured interviews with 10 faculty members involved in test design, the results showed that an OSCE can be designed for BSc or lower-level students in case of technical support and improvement, training and guidance of faculty members for the proper design of different stations and appropriate educational objectives (23). These studies, on the one hand, indicate the existence of structural and content problems in most of the stations of these tests and on the other hand, indicate the need for intervention to improve the design quality of OSCE tests. In this regard, our research showed that providing appropriate feedback with an emphasis on enhancing the strengths and eliminating the relevant weaknesses can be considered as effective ways to improve the design and implementation of these tests. One of the major drawbacks of the present study was the lack of including all dimensions of the OSCE (e.g., standardized patients and simulated patients), which should be assessed in separate studies.

Conclusion

The current framework could be used as a basis for standardization and organization of comprehensive assessment system in universities. The framework included four axes of goal, structure and organizations, design and implementation of the comprehensive assessment system, and quality assurance of the assessment system as the key axes of the assessment system. The implementation of the assessment system can facilitate to achieve "assessment for learning" approach through the organization of various components of the student assessment.

Acknowledgments

We thank the research deputy of Hamedan University of Medical Sciences for the financial support of this project. Hereby, we extend our gratitude to all professors who designed the OSCE for assisting us in performing the research.

Conflicts of Interest: The authors declare that there are no conflicts of interest.

References

1. Zimmerman B, Pons N. Student Differences in Self-regulated Learning. J ournal Educatoinal Psychology. *1990; 82(1): 51-59.*

2. Murtagh A M, Todd SA. Self-regulation a challenge to the strength model. Journal of Articles in Support of the Null Hypothesis. *2004; 3(1): 19-51.*

3. Zimmerman J, Kiysantas A. Aquiring Writing revision skill: shifting from process to outcome self – regulatory goals. Journal of Educational psychology. *1999; 91, 1-10.*

4. Zimmerman BJ, Pons MM. Development of a structured interview for assessing student use of self-regulated learning strategies. American Educational Research Journal. *1986; 23(4), 614-628.*

5. Pint rich PR, DeGroot EV. Motivational and self-regulated learning components of classroom academic performance. Journal of Educational Psychology.*1990; 82, 33-40.*

6. Linnenbrink E A, Pint rich P R. The role of self-efficacy beliefs in student engagement & learning in the classroom. Reading & Writing Quarterly. *2003; 19, 119-37.*

7. Stavrianopoulos K. Adolescents' metacognitive knowledge monitoring and academic help. College Student Journal. *2007; 41(2): 444-453.*

8. Good TL, Brophy J. Contemporary Educational Research (2nd Ed.). New York Appleton-Century-Crofts; *1995.*

9. Farokhi H, Karami A, Mir Drikvand F. The Effect of Teaching Metacognitive Strategies on Improving Academic Achievement and Reducing Test Anxiety in Nursing Students. Journal of Medical Education and Development. *2018, 13(1): 31-41.*

10. Besharat M, Abbaspour Dopplanes T. The Relationship between Metacognitive Strategies and Creativity with Resilience in Students. New Findings in Psychology. *2010; 5(14): 109-124.*

11- Groccia J E. The college success book: A whole-student approach to academic excellence. Glen bridge Publishing Ltd Viewpoints of nursing and midwifery faculty. Tabriz University of Medical Sciences. Iranian Journal of Medical Education. *(2002); 2:10-11*

12. Slavin RE. Educational psychology: theory and practice. Mohamadi Y. Tehran: Virayesh. *2006.*

13. Amini A, Valizadeh S, Mohammadi B. Survey of effective factors on learning motivation of clinical students and suggesting the appropriate methods for reinforcement the learning motivation from the viewpoints of nursing and midwifery faculty, Tabriz University of Medical Sciences. Iranian Journal of Medical Education. *2002; 2: 10-11.*

14. Salajegheh M, Hoseni Shaoon A .The role of motivation in learning. Journal of Yazd Medical Education Development Center. *2018; 13.2: 172-174.*

15. Narimani M, Khoshnoodnia Chomachaei B, Zahed A, Abolghasemi A. A compare motivational beliefs and self- regulation learning strategies in boy students with dyslexia, dysgraphia, dyscalculia and normal. Journal of Learning Disabilities. *2016; Vol. 5, No.4:141-147.*

16. Sideridis G D, Morgan P, Botsas G, Padeliadu S, Fuchs, D. Predicting LD on the basis of motivation, metacognition and psychopathology: An ROC analysis. Journal of Learning Disabilities. *2006; 39, 215-229.*

17. Tammanaeefar M, Gandomi Z. Relationship between achievement motivation and academic Achievement of students. Educational Strategies Journal. *2012; 4(1):15-19.*

18. Hossein Mardi A A, Hossein Mardi Z. Prediction of academic achievement based on emotional intelligence and achievement motivation among students of Islamic Azad University of Roudehen branch in 2013. Pajouhan Scientific Journal. *2015; 13, No 3.*

19. Solhi M, Salehfard A, Agha F, Hosseini F, Ganji M. Relationship between self-regulated strategies and creativity with the academic performance of public health students. Iranian Journal of Rah Avard Salamat: *2015; Vol. 1, No. 2,*

20. Alotaibi KN. The learning environment as a mediating variable between self-directed learning readiness and academic performance of a sample of Saudi nursing and medical emergency students. Nurse education today. *2016; 36:249-54.*

21. Agustiani H, Cahyad S, Musa M. Self-efficacy and Self-Regulated Learning as Predictors of Students Academic Performance. The Open Psychology Journal. *2016; 9.1-6.*

22. Goradia T, Bugarcic A. Can self-regulated learning strategies improve academic achievement in online higher education learning environments? Advances in Integrative Medicine. *2017; 4(1):36-37.*

23. Arami Z. Manshaee G R, Abedi, Sharifee T. the Comparison of Motivational Beliefs, Metacognitive Skills and Self-Regulation Learning between Gifted and Ordinary Students of the City of Isfahan. Journal of New

Educational Approaches.School of Educational Sciences and Psychology, University of Isfahan *2016; 11, 2. 24: 59-70.*

24. Kajbaf M B, Moulavi H, Shirazi A. R. Relationship of motivational beliefs and self-regulated learning strategies with high school students' academic performance. Advanced in Cognitive Science, *2003. 5(1), 27-33.*

25. Ostovar S. Abedi M. A Comparison of Motivational Beliefs and Self-regulating Learning Strategies among Normal and Probation Students. Journal of Educational Psychology studies. *2016; 13, 24, 1-20.*

26. Ghadampoor A, Yusefvand L, Radmehr P. The Compare ffectiveness of Cognitive and Meta Cognitive Strategies Education on Academic hope of Girl and Boy Students. Journal of Cognitive Strategies in Learning. *2018; VOL 5, NUM 9, 33-47.*

27. Pyhältö K, Soini T, Pietarinen J. Pupils'pedagogical well-being in comprehensive schoolSignificant positive and negative school experiences of Finnish ninth graders. Eur Journal Psycholog Educatoinal; *2010.25(2):207-221.*

28. Pelage O. Test anxiety, academic achievement, and self-esteem among Arab Adolescents with and without learning disabilities. Journal of Learning Disability. *2009; 32(1), 11-25.*

29. Stipek, Deborah, J. Motivation to learn: From theory to practice (4th ed.), Boston, *2002.*

30. Artino A R, La rochelle J S, Durning S J. Second-year Medical Students' Motivational Beliefs, Emotions, and Achievement. Medical Education. *2010; 44: 1203-1212.*

31. SaraeiA A, Mohammad pour M, Jajarmi M. The role of motivational beliefs on academic performance by mediating different types of goal orientations. Two Quarterly Journal of Sociology of Education. *2018; 10.134-146.*

32. SabzMakan L, KeikavoosiArani L, Hossein Sh, AliAkbarpoor S. The Relationship between Self-Regulation learning strategies, motivational beliefs and the academic performance of students At Alborz University of Medical Sciences in. Journal of Medical Educational Development. *20161; 2(3): 171-84.*

33. Bembenutty, H.The first word, a letter from the guest editor on self-regulation of learning. Journal of Advanced Academics. *2008; 20(1), 5-17.*

Biglarkhani M, Meyari A, Momtaz HE, Soltanian AR, Pezeshki N, Yavangi M, Gorbaniamjad G. The Effect of Feedback on Objective Structured Clinical Examinations: A Quasi-experimental study. J Med Educ Dev. 2021; 13 (38) :57-64