

Scoring consistency of standard patients and examiners in the developed dental OSCE system

Feng Zhu

Nanjing Stomatological Hospital, Medical School of Nanjing University

Li Wu

Nanjing Stomatological Hospital, Medical School of Nanjing University

Xiuxiu Shao

Nanjing Stomatological Hospital, Medical School of Nanjing University

Lijuan Huang

Nanjing Stomatological Hospital, Medical School of Nanjing University

Xiangfeng Meng

Nanjing Stomatological Hospital, Medical School of Nanjing University

Rongrong Nie (✉ nrr63@nju.edu.cn)

Nanjing Stomatological Hospital, Medical School of Nanjing University

Research Article

Keywords: Consistency, objective structured clinical examination, residents, standard patient, dental

Posted Date: September 19th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-2024055/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Medical Education on February 17th, 2023. See the published version at <https://doi.org/10.1186/s12909-023-04087-6>.

Abstract

Objective

This study aimed to investigate the use of standard patients (SPs) and examiners as assessors for scoring in the developed dental objective structured clinical examination (OSCE) system and to evaluate scoring differences between them.

Methods

The doctor–patient communication and clinical examination station in the OSCE examination system was developed. The examination time of this station was 10 min, and the examination institution wrote the script and recruited SPs. A total of 146 examinees receiving standardized training at the Nanjing Stomatological Hospital, Medical School of Nanjing University, from 2018 to 2021 were assessed. The same examinee was scored by SPs and examiners according to the same scoring rubrics. Subsequently, the SPSS software was used to analyze the examination results of different assessors and evaluate their correlation.

Results

The average score of all examinees at this station was 90.45 ± 3.52 by SPs and 91.53 ± 4.13 by examiners. A significant positive correlation was found between their scores, with a correlation coefficient of 0.746. Thus, the assessors had good reliability and higher consistency.

Conclusions

It was feasible to use SPs directly as assessors. They could provide a simulated and realistic clinical setting and create favorable conditions for comprehensive competence training and improvement for medical students during school.

Introduction

Objective structured clinical examination (OSCE), a multi-station clinical skill examination, is an approach to assessing the clinical competence of medical students ^[1, 2]. This novel clinical competence assessment mainly uses a series of pre-designed simulated clinical settings to evaluate the clinical competence of medical students; the examinees complete the tasks designed at each station and are assessed in multiple stations simulating clinical settings. Since the 1990s, OSCE has been included in the curricula of several dental schools worldwide to assess the competence of medical students in various parameters, including communication, patient education, clinical skills, and critical thinking ^[3, 4].

A doctor–patient communication and clinical examination station, which is an important station in the OSCE examination, performs standardized and systematic training for healthy individuals to make them become standard patients (SPs). Further, SPs can accurately present actual clinical problems and imitate the symptoms of the corresponding cases, including body movements, pain degree, facial expressions, self-report of symptoms in the medical history, and so forth. Examinees understand the medical history and conditions according to the symptoms of SP to make a correct diagnosis. In the examination, SP is a simulated patient who “undergoes medical history and physical examinations,” and a well-trained SP can act as an examiner to assess the performance of examinees and even guide them. Thus, SP provides the opportunity for medical students to practice communication skills and obtain patients’ feedback, which is the most valuable and unobtainable in real life [5].

Compared with other assessment methods, the “objectification” and “structuration” of OSCE can well simulate the real clinical settings and achieve consistency in assessment contents. However, OSCE requires numerous examiners and examination persons during the actual implementation to develop detailed pre-exam plans. The limitation of examination time and space and the lack of relevant examination persons are the common obstacles reported in the literature that affect the promotion and implementation of OSCE [6, 7]. Some studies compared the scoring differences of different assessors (full-time teachers, part-time teachers, residents, or examinees) on the examination results of examinees in the doctor–patient communication and clinical examination station [8–10]. However, the data regarding the use of an SP as an examiner for scoring are few.

This setting was used to construct a suitable doctor–patient communication and clinical examination station, improving the rating scale and pre-exam training. The SP can undertake the station evaluation work and give feedback and guidance to the examinees, which is closer to clinical practice and can reduce the large numbers of examiners in the OSCE examination. In the present study, the scoring differences between SPs and professional examiners were assessed based on a 4-year resident completion evaluation.

Methods

Participants and Methods

The present study was approved by the Institutional Review Board of Nanjing Stomatological Hospital, Medical School of Nanjing University (No. NJSH-2022NL-074). The OSCE examination system was used to assess 146 residents in the hospital from 2018 to 2021. During the examination, the same examinee was scored by an SP and an examiner according to the same scoring rubrics, and brief feedback was written on the rating scale for each examinee. The study was conducted in strict accordance with the requirements of the Declaration of Helsinki. Participation was voluntary, without any compensation or incentive. We guaranteed both confidentiality and anonymity, and participating residents completed an informed consent form.

Construction of SP station

The OSCE had eight stations, of which the third station was the doctor–patient communication and clinical examination station. It used an SP to assess the students' doctor–patient communication and clinical examination competencies and whether students could perform standardized receptions. The SP was also used to comprehensively and accurately collect the medical history based on chief complaints, make a possible diagnosis, and determine further examination and treatment plans in combination with the clinical examination results. Meanwhile, the students' communication skills, such as appearance, attitude, and language expression during clinical reception, were assessed, including inquiry for patients, diagnosis and treatment decisions, disease prognosis, diagnosis and treatment costs, and so forth (Table 1). In this station, the examination time was 10 min, and the score was calculated using the centesimal system, accounting for 8% of the total score^[11].

Recruitment of SPs

Based on the actual conditions of dental teaching, four scripts were written: patients with acute pulpitis, patients with pericoronitis of wisdom teeth, patients with gingivitis, and patients with tooth defects; these dental diseases were quite common. One week before the examination, oral examinations were performed on the voluntary participants at this hospital. Those meeting the script requirements were included, and SPs were preferentially selected according to their own wishes. Two SPs were recruited for each script, totaling eight SPs. The rotation was made every 2 h during the examination.

Pre-exam training

After SP recruitment, teachers with SP training experience distributed and explained their scripts and watched real clinical patient videos. SPs and station examiners were included 3 days before the examination. Four examiners were teachers with more than 3 years of teaching experience. Each examiner corresponded to two SPs of the same script, which were assigned to four groups. In addition, the chief examiner in charge of the SP station explained the scoring rubrics to SPs and examiners according to the rating scale, reviewed the assessment videos of the previous year, and conducted simulated scoring. At the same time, eight simulated examinees performed field training, and SPs performed and scored with examiners. The passing criteria were clarified again.

Statistical analysis

Statistical analyses of the examination results were performed using SPSS 26.0 software. The independent-samples *t* test was performed for the scores of SPs and examiners on each item. Spearman's rank correlation test was used for the correlation between them.

Table 1
Rating scale of the doctor–patient communication and clinical examination station

Item		Requirements	Score	Points
Doctor–patient communication (50)	Professional behavior	Wear neatly and standardly and meet professional requirements	10	
	Personal emotion	Various behaviors when interacting with patients, including gestures, facial expressions, or confidence during a conversation	10	
	Language communication	Able to use simple language to explain patient's problems and ask about relevant medical history for chief complaints during communication	10	
	Relationship building	Build a harmonious relationship with the patient and express comfort and care about the patient's disease experience so that the patient is happy to communicate	10	
	Patient management	Provide hygiene guidance to patients and propose further treatment plans and recommendations	10	
Clinical examination (50)	Preparation before examination	Proper hand washing, wearing gloves, and adjusting the chair position to keep the patient in a comfortable position	10	
	Communication during the examination	Inform the patient before the examination and pay attention to the patient's feelings during the examination	10	
	Examination methods	Select appropriate examination instruments and perform examinations in a correct manner	10	
	consciousness of patient-friendly and dedication	Be careful not to cause pain and discomfort to the patient during the examination	10	
	End of examination	Take off gloves and wash hands at the end of the examination; keep the chair in place to facilitate the patient to leave	10	

Results

Table 2 shows that the independent-samples *t* test was used to examine the scoring differences between SPs and examiners. The results suggested that SPs and examiners did not differ significantly in terms of annual scores and 4-year total scores on examinees in the doctor–patient communication and clinical examination station in the standardized training completion examination from 2018 to 2021 ($P > 0.05$). Moreover, Spearman's rank correlation test was used to investigate the consistency of examiners and SPs

in each item scoring, namely the reliability of assessors. The results showed a significant positive correlation in annual scores and 4-year total scores between SPs and examiners ($P < 0.001$) (Table 3). In addition, Table 4 analyzes the 10 scoring items of this station. A significant positive correlation was observed between SPs and examiners in the scores. The reliability of assessors was good, and the consistency was higher ($P < 0.001$).

Table 2
Scores of SPs and examiners

Year	Number of examinees	SP Mean \pm SE	Examiners Mean \pm SE	<i>t</i>	<i>p</i>
2018	36	89.86 \pm 3.63	90.58 \pm 4.59	-0.740	0.462
2019	37	89.05 \pm 3.54	90.70 \pm 4.63	-1.722	0.089
2020	34	92.56 \pm 2.72	92.94 \pm 3.06	-0.544	0.588
2021	39	90.46 \pm 3.25	91.97 \pm 3.70	-1.919	0.059
Overall	146	90.45 \pm 3.52	91.53 \pm 4.13	-1.426	0.137

Table 3
Correlation analysis of scores of SPs and examiners

Year	Number of examinees	SP	Examiners	Spearman's rho	<i>P</i>
2018	36	89.86 \pm 3.63	90.58 \pm 4.59	0.749**	0.000
2019	37	89.05 \pm 3.54	90.70 \pm 4.63	0.851**	0.000
2020	34	92.56 \pm 2.72	92.94 \pm 3.06	0.662**	0.000
2021	39	90.46 \pm 3.25	91.97 \pm 3.70	0.667**	0.000
Overall	146	90.45 \pm 3.52	91.53 \pm 4.13	0.746**	0.000
**Represents $P < 0.001$.					

Table 4
Correlation of scores of SPs and examiners on 10 items

Item	<i>n</i>	Examiners	SP	Spearman's rho	<i>P</i>
Professional behavior	146	9.70 ± 0.65	9.66 ± 0.70	0.651**	0.000
Personal emotion	146	9.43 ± 0.67	9.48 ± 0.76	0.412**	0.000
Language communication	146	8.31 ± 1.07	8.48 ± 1.46	0.442**	0.000
Relationship building	146	9.58 ± 0.82	9.65 ± 0.80	0.439**	0.000
Patient management	146	7.79 ± 1.36	7.99 ± 1.41	0.951**	0.000
Preparation before examination	146	9.07 ± 0.80	9.27 ± 0.89	0.369**	0.000
Communication during the examination	146	9.08 ± 0.93	9.20 ± 0.94	0.639**	0.000
Examination methods	146	9.34 ± 0.93	9.36 ± 0.97	0.533**	0.000
Consciousness of patient-friendly and dedication	146	9.10 ± 0.76	9.24 ± 0.83	0.382**	0.000
End of examination	146	9.06 ± 0.73	9.20 ± 0.71	0.410**	0.000
**Represents $P < 0.001$.					

Discussion

The widespread application of SPs in medical education reflects the improvement in humanistic quality and care in society. Through trained SPs, students who are going to work are taught how to communicate with patients and deal with unexpected problems, experience humanistic care for patients, and improve their abilities to identify, analyze, and deal with problems. The clinical thinking abilities of medical students can be improved by SP inquiry and physical examinations, communicating with SPs, understanding patients' experience of illness, and providing health education to patients^[12]. Through personal experience combined with the actual situation and scoring rubrics, SPs assess students, record and identify their shortcomings, and provide students with a realistic and comprehensive clinical process and real feelings. Applying SPs under the framework of OSCE makes the assessment more fair and just. Each SP is trained for one aspect, so that each SP faces the same patients and the same problems,

making the evaluation fair and true and thus avoiding previous biases caused by collecting medical histories and signs of different patients by different students.

OSCE is designed to standardize the examination and reduce variables that may affect performance assessment. Thus, for a well-designed OSCE, the examination results of examinees are mainly affected by their own competencies, ensuring minimal interference from other variances. The consistency of SP performance for each examinee is crucial for the SP station. Poorly standardized SPs perform differently for different examinees, reducing the reliability of examination^[13]. Therefore, the training of examiners and SPs is an important element in the quality assurance and standardization process before the examination^[14]. In the present study, SPs were trained three times before the examination, and SPs and matched examiners were trained two times. The training not only included the explanation of scripts and scoring rubrics but also analyzed previous examination videos and organized the training of simulated examinees. A checklist with detailed contents was designed, listing the assessment points of examinees in communication, to increase the consistency of the assessment. Through careful pre-exam design and training, regardless of examiners, these measures can reduce the differences in examiners and improve behavioral consistency, thereby increasing the reliability of the examination^[15]. The results showed no significant difference in the annual scores between SPs and examiners. In the four years from 2018 to 2021, a significant positive correlation was found between SPs and examiners in the scores of 10 items at this station, suggesting good reliability and higher consistency of assessors. However, some studies demonstrated statistical differences in the scores given by different types of examiners to the same examinees; this result might be associated with their insufficient preparation for the examination^[16].

The selection of SP as the station examiner can reduce large numbers of examiners at OSCE. More importantly, real feedback can be obtained from patients. A previous study has also shown that medical students believe that the most valuable part of SP training and assessment is the direct feedback from patients, "This is a very rare opportunity because we rarely obtain this feedback from real patients and their families"^[17]. During the assessment, it was also found that the SP was more likely to propose some details that might be neglected by the examiners, such as "he (she) doesn't look at me during the communication," "my speech is interrupted," "he (she) is unable to empathize with my pain," "the movement during the examination is rude," "the explanation of the conditions is too professional," and so forth. As professionals, examiners may be familiar with disease settings and think that comfort and explanation are unnecessary. At the same time, as a bystander, examiners cannot share the experience of the patients gained by the manipulation of the examinees. Thus, a qualified SP understands case contents well and can provide positive, helpful feedback on the performance of examinees.

The type of patients required for a SP station depends on assessment contents and the role they are expected to play. The assessment contents in the present station require examinees to perform an oral examination based on the chief complaints, provide possible diagnosis and treatment plans based on the examination results, and communicate with patients. Given oral diseases are often obvious, for humanitarian reasons, the oral disease status of the SP cannot be maintained without treatment for a

long time. Therefore, SPs are selected and trained before the assessment, and they cannot perform long-term tasks. Moreover, the replacement frequency is high. Thus, experienced training teachers and a whole set of standardized selection and training processes are required to ensure the consistency of the examination. Other limitations of the present study include small sample size of students and examiners and a shorter observation time so that the conclusions are not generalized. More SPs will be used as assessors and instructors in the education of medical students in the future.

Conclusions

The present study demonstrated that the scoring consistency between SPs and examiners was excellent, suggesting that it was feasible to use SPs directly as assessors after pre-exam comprehensive and detailed training and a well-designed rating scale. They can be used for teaching assessment and usual teaching and training, as well as to provide simulated and realistic clinical settings, thus creating favorable conditions for the comprehensive competence training and improvement in medical students during school.

Declarations

Ethics approval and consent to participate

This study was approved by the Institutional Review Board of Nanjing stomatological Hospital, Medical School of Nanjing University (No. NJSH-2022NL-074).

The study was conducted in strict accordance with the requirements of the Declaration of Helsinki. Participation was voluntary, without any compensation or incentive. We guaranteed both confidentiality and anonymity, and participating residents completed an informed consent form.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study is funded by the “3456” Cultivation Program For Junior Talents of Nanjing Stomatological School, Medical School of Nanjing University 0222E301 .

Authors' contributions

Author Zhu Feng is the lead author of the paper.

Author Wu Li is the main implementer of the experiment and the main administrator of the OSCE examination.

The author Shao Xiuxiu was responsible for the collection, collation and statistical analysis of the experimental data.

Authors Lijuan Huang and Xiangfeng Meng were the main designers and examiners of the experiment.

Author Nie Rongrong was responsible for the overall framework and review of the article.

Acknowledgements

The author thanks Prof. Weibin Sun, Sijing Xie, the SPs and staff of Education Department, Nanjing Stomatological Hospital; and last but not least, the students participant in the study.

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 Sep;35(9):e1437-46.
2. Brannick MT, Erol-Korkmaz, et al. A systematic review of the reliability of objective structured clinical examination scores ^[J]. *Med Educ*. 2011;45(12):1181–9.
3. Graham R, Bitzer LAZ, Mensah FM, et al. Dental student perceptions of the educational value of a comprehensive, multidisciplinary OSCE. *J Dent Educ*. 2014 May;78(5)(1):694–702.
4. Park SE, Kim A, Kristiansen J, et al. The influence of examiner type on dental students' OSCE scores. *J Dent Educ*. 2015 Jan 1;79(1):89–94.
5. Ream M, Albert DVF, Lash T, Verbeck N, Weisleder P. Improving Child Neurology Residents' Communication Skills Through Objective Structured Clinical Exams. *MedEdPORTAL*. 2021 Mar 4;17:11120.
6. Navickis MA, Bray KK, Overman PR, et al. Examining clinical assessment practices in U.S. Dental Hygiene Programs. *J Dent Educ*. 2010 Mar 1;74(3):297–310.
7. Solà M, Pulpón AM, Morin V, et al. Towards the implementation of OSCE in undergraduate nursing curriculum: a qualitative study. *Nurse Educ Today*. 2017 Feb;49:163–7.
8. Park SE, Kim A, Kristiansen J, Karimbux NY. The influence of examiner type on dental students' OSCE scores. *J Dent Educ*. 2015 Jan;79(1):89–94.

9. de la Cruz MS, Kopec MT, Wimsatt LA. Resident Perceptions of Giving and Receiving Peer-to-Peer Feedback. *J Grad Med Educ*. 2015 Jun;7(2):208–13.
10. Andrews E, Dickter DN, Stielstra S, Pape G, Aston SJ. Comparison of Dental Students' Perceived Value of Faculty vs. Peer Feedback on Non-Technical Clinical Competency Assessments. *J Dent Educ*. 2019 May;83(5):536–45.
11. Rongrong Nie F, Zhu X, Meng, et al. Application of OSCE for Stage Assessment in Standardized Training for Oral Residents. *J Dent Educ*. 2018 Sep;82(9):1000–6.
12. Bosse HM, Schultz JH, Nickel M, et al. The effect of using standardized patients or peer role play on ratings of undergraduate communication training: a randomized controlled trial. *Patient Educ Couns*. 2012;87:300–6.
13. Smee S. ABC of learning and teaching in medicine 'Skill based assessment'. *BMJ*. 2003;326:703–6.
14. Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part II: organisation & administration. *Med Teach*. 2013 Sep;35(9):e1447-63.
15. Tan CP, Azila NM. Improving OSCE examiner skills in a Malaysian setting. *Med Educ*. 2007;41:517.
16. Park SE, Kim A, Kristiansen J, Karimbux NY. The influence of examiner type on dental students' OSCE scores. *J Dent Educ*. 2015 Jan;79(1):89–94.
17. Chen TC, Lin MC, Chiang YC, Monrouxe L, Chien SJ. Remote and onsite scoring of OSCEs using generalisability theory: a three-year cohort study. *Med Teach*. 2019;41(5):578–83.