

# Application of OSCE for Stage Assessment in Standardized Training for Oral Residents

Rongrong Nie, Feng Zhu, Xiangfeng Meng, Hai Zhang, Sijing Xie, Li Wu, Jie Yang, Weibin Sun

**Abstract:** The aim of this study was to establish and implement an objective structured clinical examination (OSCE) for stage assessment in standardized training for oral residents, evaluating its validity and suitability. An OSCE was established for stage assessment of 158 residents in a program in Jiangsu Province, China. Its validity and suitability were evaluated using indices such as reliability and discrimination coefficients of assessment results. The established OSCE had eight stations: interpretation of auxiliary examination results, basic knowledge, doctor-patient communication and clinical examination, clinical case analysis, medical record and prescription writing, public skills, first aid skills, and specialist's operation skills. The mean overall score and consistency coefficient (Cronbach's alpha) for the test subjects were  $83.64 \pm 3.69$  points and 0.732, respectively. The developed OSCE was reasonably established, with credible assessment results, and indices such as mean and discrimination coefficient of test scores from each station were appropriate. Therefore, the proposed protocol was found to be reliable and suitable.

Rongrong Nie, BDS, PhD, is Associate Professor, Department of Prosthodontics, Nanjing Stomatological Hospital, Medical School of Nanjing University, China; Feng Zhu, MD, PhD, is Associate Professor, Department of Oral and Maxillofacial Surgery, Nanjing Stomatological Hospital, Medical School of Nanjing University, China; Xiangfeng Meng, BDS, PhD, is Associate Professor, Department of Prosthodontics, Nanjing Stomatological Hospital, Medical School of Nanjing University, China; Hai Zhang, DMD, PhD, is Associate Professor, Department of Restorative Dentistry, University of Washington; Sijing Xie, BDS, PhD, is Deputy Chief, Department of Teacher Education, Nanjing Stomatological Hospital, Medical School of Nanjing University, China; Li Wu, BDS, MS, and Jie Yang, BDS, PhD, are the main examination organizers, Department of Teacher Education, Nanjing Stomatological Hospital, Medical School of Nanjing University, China; and Weibin Sun, BDS, PhD, is Professor and Deputy Dean (Teaching), Nanjing Stomatological Hospital, Medical School of Nanjing University, China. Direct correspondence to Prof. Weibin Sun, Department of Teacher Education, Nanjing Stomatological Hospital, Nanjing University, Nanjing, China; +86-25-8362-0127; wbsun@nju.edu.cn.

**Keywords:** advanced dental education, oral residents, assessment, clinical competence, objective structured clinical examination, standardized training

*Submitted for publication 8/7/17; accepted 1/23/18*  
*doi: 10.21815/JDE.018.099*

Standardized training for residents is an important part of dental education following graduation. In lifetime dental education, it serves as a predominant link between basic training in dental school and continuing education. Therefore, it is extremely important in obtaining high-level dentist training and improving the quality of dental care. Developing regulations concerning standardized training for residents has been initiated in China since 2014. Therefore, supporting assessment systems are needed to evaluate assessment results and determine whether residents meet the training requirements.

Traditional assessments focus mainly on evaluating a participant's knowledge; however, in real-world practice a competency-based assessment is preferred.<sup>1</sup> In the practice of dental education, we have also found that outstanding issues in training and assessment for clinical competence are non-rationality, arbitrariness, and subjectivity, which result in the inability to provide an accurate, objective, and comprehensive assessment of the medical student. Moreover, accurate objectives and grading

scales for assessing operational skills of students are lacking. Clinical operational assessment for students is usually a clinical admission test in which the student is confronted with a randomly selected patient, requiring him or her to provide relevant treatments accordingly.

Currently, no contents or grading scales tailored for students at each specific stage exist, and assessment results are largely determined by reviews arbitrarily. Therefore, different oral conditions, operation difficulties, and communication abilities all contribute to the inability to genuinely reflect the actual clinical competence of the student. Meanwhile, the low credibility and incomparability of clinical assessment results make clinical assessment a mere formality. An objective structured clinical examination (OSCE), also known as a multi-station clinical skills examination, is a test format to evaluate the clinical competence of health professions students. The most distinguishing characteristic as well as advantage of this test is the possibility of direct and objective assessment of clinical skills of participants. In addition,

it is widely applied in medical licensing examination and residency training testing in Western countries.<sup>2-5</sup> The OSCE is promising in the future as a standardized test focusing on competency assessment. Compared with traditional tests, more multi-evaluation items (i.e., multiple testing stations and contents) in the OSCE have enhanced its authority, providing a more comprehensive and direct manifestation of the value of this test. Meanwhile, it also avoids paradoxes between rigorous medical practices and strict clinical examinations. Therefore, OSCE complies with the trend of international medical education and assessment.<sup>6</sup>

OSCE is a global health professions education method used for assessing clinical competence in the last 30 years. It was originally proposed by Harden and Gleason from the University of Dundee in 1975, and first reported as a novel test module for assessing clinical skills of medical students in 1979.<sup>7</sup> This novel module assesses mainly the clinical competence of students via a series of predesigned simulated clinical scenarios, during which participants are required to complete all the specified tasks in various scenarios and are graded accordingly. The basic rationale for OSCE is performance-based testing, which provides an objective, ordered, and organized assessment framework that assesses clinical competence. It is also an assessment module that evaluates knowledge, skills, and attitudes, constituting a good platform for a comprehensive assessment of the clinical competence of residents.<sup>8,9</sup> The aim of this study was to establish and implement an OSCE for stage assessment in standardized training for oral residents, evaluating its validity and suitability.

## Methods

The project was approved by the Institutional Review Board of Nanjing Stomatological Hospital of Nanjing University (2016NL-078). In June 2016, a stage assessment was conducted among 158 dental residents (66 males and 92 females) enrolled in a program in Jiangsu Province, China.

The OSCE is not a specific assessment method, but it offers an objective, orderly, and organized frame of assessment, within which each institution can add unique contents and methods for assessment based on its own courses and examination syllabi. An OSCE for dental residents was developed in this study referring to the OSCE model and integrating the current requirements of advanced dental education, national practice physician examination, and standardized training for residents. This assessment program reflected the actual clinical competence of participants, using objective and structured components.

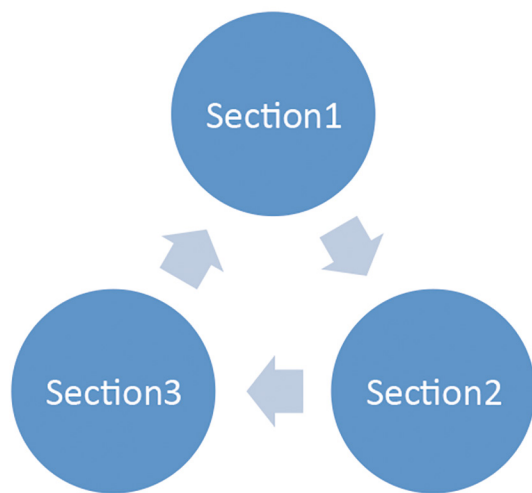
This assessment was a multi-station grading system comprising three sections of eight stations each. Standardized patients (SPs), simulated humans, standard head models, and dental bionic teeth were used for assessing all clinical operations to ensure assessment fairness. Each participant was required to occupy all eight stations, each for five to 30 minutes, and complete the tasks at each station within the time specified (Table 1). When the examination was carried out, each of three total examination sections included eight students. The distribution of candidates and examiners in each section is shown in Table 2.

**Table 1. Contents, forms, and durations of each objective structured clinical examination in assessment section**

Station	Item	Content	Duration (min)	Weighted Total Score (%)	
Section one	I	Auxiliary exam results interpretation	20 multiple choice items (oral X-rays)	20	17
	II	Basic knowledge	25 multiple choice items	20	17
Section two	III	Doctor-patient communication and clinical examination	Apply oral exams to standardized patient and collect patient history	10	8
	IV	Clinical case analysis	Analyze case and answer questions from the reviewer based on standardized case report	20	17
	V	Medical record and prescription writing	Write a medical record and prescription for narcotic drugs based on standardized case report	10	8
Section three	VI	Public skills	Measure blood pressure of each other	5	4
	VII	First-aid skills	Apply cardiopulmonary resuscitation to the standardized simulated human	5	4
	VIII	Specialist skills operation	Apply performances related to specialist skills to the standardized head model	30	25

**Table 2. Distribution of students and examiners in each station**

Station		Number of Students	Number of Interviewers	Duration (min)
Section one	I	4	2	20
	II	4		20
Section two	III+V	4	8	10+10
	IV	4	8	20
Section three	VI	8	16	5
	VII			5
	VIII			30



**Figure 1. The rotation sequence of the stations**

Stations I and II used a computer to answer objective questions, and two examiners were responsible for this. Every other assessment of each station was performed by two examiners. The time in each section was 40 minutes with a sequence of Section 1 to Section 2 to Section 3 to Section 1 (Figure 1). With complete rotations, 24 students could be examined in two hours. A total of 158 candidates were divided into seven groups, with two-hour intervals between each group. Two days in total were needed.

Although no patients or head models were involved in stations I, II, IV, and V, the tasks, which assessed the skills of interpreting examination results and providing a diagnosis as well as suggestions relevant to treatment, were diverse. Usually, the human-computer-interaction method was adopted. The participants were asked to answer the questions directly on a computer, which provided an assessment score.

The way they answered the questions did not matter in this process, and a score was assigned based only on the answers provided. Station III, termed patient communication and clinical examinations, assessed whether a participant could admit an SP in an acceptable way, collect the patient history according to the chief complaint, and provide possible diagnoses and advice for further examinations and therapeutic plans based on his or her own clinical examination. In addition, this station evaluated communication skills, including appearance, attitude, and language use, during admission and inquiry, as well as the skills for questioning the patient and providing information regarding the diagnostic and therapeutic decisions, disease prognoses, and medical costs. Components of operational skill assessment included stations VI, VII, and VIII, which assessed the mastery of public skills, first-aid skills, and specialist skills, respectively. A participant was required to go through all these processes before assessment by viewers based on the standardized grading scale.

The assessment was organized by the Department of Education of the Affiliated Stomatological Hospital of Medical School of Nanjing University and conducted at the Clinical Simulation Center of Nanjing University. All evaluators at various discipline stations were university teaching staff for the respective disciplines. The assessment criteria were printed as a written checklist of eight to 15 predetermined items (Table 3). These criteria were set by the respective clinical disciplines and were based on clinical assessment rubrics of the Institute of Dentistry. A 100%-based system was used in each station, and the final score was the sum of weighted scores from all stations. The OSCE was used as a summative assessment method, with the gateway examination employed to complete residency training.

Preparation meetings for SPs, participants, and instructors were held prior to the examination. SPs were trained for job responsibilities, assigned tasks, and provided with targeted script training as well as feedback-based real-world simulation practice. Examination contents and forms, as well as pre-examination frequently asked questions, were explained to the participants who visited the test places. Grading criteria were explained to the examination instructors in each department, and pass criteria were established.

The reliability of an OSCE indicated the extent to which examination scores were reproducible. The reliability of the proposed OSCE was evaluated using

**Table 3. Item for the prosthodontics-grading scale for operation in anterior porcelain-fused-to-metal (PFM) crown preparation**

Principle	Feature	Ideal	Unacceptable
Structural durability	Incisal reduction	2.0 mm	>2.2 mm or <1.8 mm
	Axial reduction	1.2 mm (labial) 0.5 mm (MDL)	>1.4 mm or <1.0 mm >1.4 mm or <1.0 mm
Retention resistance form	Lingual concavity reduction	1.0 mm	>0.9 mm or <0.3 mm
	Taper	6°	>9° or <3°
	Length	>3 mm	Short, not retentive
Marginal integrity	Draw	Parallel to long axis	Undercut, mal-aligned
	Finish line	1 mm above soft tissue	>1.5 mm or <0.5 mm or irregular
	Labial chamfer	100° CSA	>110° or <90°
	Lingual chamfer	100° CSA	>110° or <90°
Convenience	Gingivoaxial line angle	Slightly rounded	Sharp
	Line angles	Rounded to fingertip	Sharp, rough, irregular
Conservation	Finish	Fine diamond	Coarse diamond, polished
	Neatness	Clean, dry, presentable	Not clean, not presentable
	Adjacent teeth	No visible marks	Marks, alterations

Cronbach's alpha, as in a similar study.<sup>10</sup> Cronbach's alpha is a measure of internal consistency, that is, how closely a set of items are related as a group. Discrimination is a statistical parameter that reflects the ability of each station to distinguish actual competences among participants. In a discrimination analysis, the total score usually reflects the actual competence of a participant, while the correlation coefficient between a station and the total score is considered the discrimination of the given station. The Spearman's rank correlation analysis was used for objective questions in stations I and II. The Pearson's correlation analysis was used for subjective questions, which were regarded as non-equidistant continuous variables with a sample size more than 30.

## Results

A total of 158 participants attended the stage assessment in the 2016 standardized training programs for dental residents in Jiangsu Province. There

were eight examination stations. Individual station and overall test scores were assessed for normality. For normally distributed variables, a p-value <0.05 indicated a statistically significant difference. The results indicated that all test scores from various stations and overall scores were normally distributed (Table 4).

The reliability coefficient (Cronbach's alpha) for internal consistency between OSCE station and overall scores was 0.732. Table 5 shows correlation coefficients between the individual station and overall scores, as well as Cronbach's alpha values after deleting a given station.

## Discussion

The traditional clinical examination comprised several variables: the student, the patient, and the examiner. In the structured clinical examination, two variables (the patient and the examiner) were more controlled, and a more objective assessment

**Table 4. Normality test and descriptive analysis for individual station and overall test scores**

Station	Minimum	Maximum	Mean	SD	p-value
I	67.00	96.00	81.18	5.91	0.104
II	64.00	96.00	80.20	6.30	0.051
III	73.00	96.00	86.20	4.48	0.494
IV	70.00	97.00	85.03	5.40	0.249
V	70.00	98.00	86.47	6.16	0.113
VI	70.00	97.00	87.60	5.20	0.084
VII	72.00	95.00	85.11	3.53	0.180
VIII	60.00	96.00	84.13	7.09	0.130
Overall score	75.49	92.00	83.64	3.69	0.666



**Table 5. Results of each station score**

Station	Correlation Coefficient Between Each Station Score and Overall Score	Cronbach's alpha After Deleting the Station
I	0.642	0.678
II	0.666	0.679
III	0.504	0.696
IV	0.531	0.702
V	0.540	0.686
VI	0.481	0.685
VII	0.348	0.709
VIII	0.702	0.724

of student's clinical competence was made.<sup>11</sup> In the proposed OSCE assessment system, the patient factor was standardized by a standardized patient, standardized head model, and artificial teeth. Each examinee faced a consistent oral environment, decreasing assessment error caused by different patients. The clinical assessment system was created by the Institute of Dentistry and based on the clinical assessment rubric that defined performance criteria for the most important skills in clinical dentistry. However, it caused a problem. All the examinees faced a standardized head model, artificial teeth, and consistent occlusal relationship, which could not truly simulate the various oral conditions in the clinic.<sup>12</sup> Thus, except for assessment on operation, some clinical cases were added in the OSCE. The stimulation on clinical actual situations was performed through the station III doctor-patient communication and station IV clinical case analysis. The clinical case analyses assessed the comprehensive analytical ability by providing the patient's chief complaint, as well as results from clinical and special examinations. The examiners asked the examinees three or four independent questions based on the case report at hand. Moreover, the physiopathological data of the patient were provided to the examinees for the further assessment of comprehensive analytical ability to better simulate the real-world clinical practice with disease evolution.

Traditional health professions education focuses on teaching professional knowledge and skills to students and neglects the fostering of humanistic qualities and medical ethics, inevitably resulting in weak awareness of service, decreased social adaptability, and inability to satisfy the multi-level social demands regarding the contents of medical services, effectiveness, and attitudes. Lack of medical humanistic care and decreased trust between physicians

and patients trigger tensions. Other studies have used OSCEs to evaluate brief communication skills training for dental students. The OSCE is a unique evaluation tool that can be used to provide a standardized assessment of student skills, medical history recording, and treatment planning.<sup>11,12</sup> Thus, a station for doctor-patient communication was specifically designed in this study, and situations of acute pulpitis, opsigenes pericoronitis, gingivitis, tooth defects, and other common diseases based on SP conditions were simulated. The SPs would simulate typical symptoms and signs, with students required to collect patient history, complete specialist examinations, provide diagnoses, develop treatment plans, and conduct preoperative communication. The students were graded by the SPs based on their observations and experiences.

For examiner factor, the selected examiners were teachers with years of teaching experience. Besides, the scoring was not based on the experience and expression. Each operation test had a detailed score table and a deduction item. Meanwhile, step-by-step operation specification was provided to the examiners for reference. Each operation of the examinee was scored by two examiners to make the test more fair and impartial and eliminate the possible partiality and shielding between examinees and examiners.

The application of an electronic information platform was one of the advantages of the proposed system. Before the examination, the questions were input into the system, including single-choice questions of stations I and II, clinical case analysis of station IV, and specialist skills operation of station VIII. After the examinees logged into the system, the questions were randomly selected and grouped from the question bank. The operation item was randomly drawn, and the examiners were randomly assigned after the examinees entered the examination room, effectively avoiding the possible leaking of questions and examiner shielding. Meanwhile, the electronic system was quite efficient for rapid correction and timely feedback of objective single-choice question after the examination and decreased the correction error rate down to the lowest. However, a stable network was needed for the successful use of electronic information system during the examination, and electronic equipment maintenance personnel was also necessary.

The main disadvantage was the increased preparation required. OSCE required many examiners to be present, all with appropriate professional training; an appropriate testing location with a room

for multiple stations; and time for preparation of the circuit, including preparation of models, radiographs, and charting. A total of 34 examiners and a chief examiner, five examination staff, and eight SPs were used for this test. In addition to full preparation, the work of examiners and staff was also needed to complete this test for so many students in two days. Despite attempts to simulate the clinical condition, a difference still existed in comparison with actual operation. For example, the SP could not be used for a truly invasive operation, and students must complete the assessment during a certain time period. Thus, the level of knowledge and skills needs to be assessed not only by the test results but also in combination with their long-term clinical performance.

Reliability is an important index, in addition to contents, forms, and maneuverability, for applying and promoting a novel and advanced teaching evaluation technique. In dental education, the reliability of OSCEs varies widely, with coefficient alpha values between 0.40 and 0.91.<sup>12-14</sup> Values of less than 0.4 indicate poor agreement, whereas the values 0.4-0.8 and >0.8 indicate fair-to-good and excellent agreements, respectively.<sup>15</sup> Correlation coefficients between individual station scores and the overall score reflect associations of the score obtained for various stations with the overall score. It is generally believed that a coefficient >0.20 indicates a strong correlation. Correlation coefficients between various station scores and the overall score were between 0.348 and 0.702, indicating that associations of individual station scores with the overall score were strong, hence appropriately achieving the goal of OSCE.

Next, Cronbach's alpha was calculated after deleting a given station. The increased value after such a deletion indicated that omitting the given station could raise the overall examination reliability. Conversely, a reduction indicated reduced overall examination reliability after omitting the given station. The OSCE in this study had a high reliability as measured by Cronbach's alpha (0.732). The high reliability indicated that the predetermined assessment criteria were in accordance with the competencies and core knowledge of the curriculum. The examination included eight stations and almost 50 questions that evaluated eight subject areas. A reasonably large number of stations and questions and task similarity among the stations likely contributed to high reliability of the examination. Interestingly, Cronbach's alpha values after deleting each station, in turn, were

between 0.678 and 0.724, indicating that deleting any station would lead to reduced overall examination reliability. These results demonstrate a suitable design for each station in the proposed OSCE system.

---

## Conclusion

The OSCE examination system could be used to evaluate clinical operation skills, communication skills, and comprehensive thinking ability to a certain degree after standardized training for oral resident physicians. The whole examination was smooth, and the statistical result indicated that the test station setting and examination content limit were reasonable and effective. The results suggested that the OSCE could be used as a tool for assessing the efficient use of teaching resources that could be educationally beneficial.

---

## Acknowledgments

Financial support for this project was provided by the Health and Family Planning Commission of Jiangsu Province (C2016035).

---

## REFERENCES

1. Schuwirth L, Ash J. Assessing tomorrow's learners: in competency-based education only a radically different holistic method of assessment will work—six things we could forget. *Med Teach* 2013;35:555-9.
2. Cooke M, Irby DM, Sullivan W, Ludmerer KM. American medical education 100 years after the Flexner report. *N Engl J Med* 2006;355(13):1339-44.
3. Lukas RV, Adesoye T, Smith S, et al. Student assessment by objective structured examination in a neurology clerkship. *Neurology* 2012;79(7):681-5.
4. Yu TC, Wheeler BR, Hill AG. Effectiveness of standardized clerkship teaching across multiple sites. *J Surg Res* 2011;168(1):e17-23.
5. Brannick MT, Erol-Korkmaz HT, Prewett M. A systematic review of the reliability of objective structured clinical examination scores. *Med Educ* 2011;45(12):1181-9.
6. Bergus GR, Kreiter CD. The reliability of summative judgements based on objective structured clinical examination cases distributed across the clinical year. *Med Educ* 2007;41(7):661-6.
7. Harden RM, Gleeson FA. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ* 1979;13(1):39-54.
8. Höfer SH, Schuebel F, Sader R, Landes C. Development and implementation of an objective structured clinical examination (OSCE) in CMF-surgery for dental students. *J Craniomaxillofac Surg* 2013;41(5):412-6.
9. Landes CA, Hofer S, Schuebel F, et al. Long-term prospective teaching effectivity of practical skills training and a first OSCE in craniomaxillofacial surgery for dental students. *J Craniomaxillofac Surg* 2014;42(5):e97-104.

10. Brown G, Manogue M, Martin M. The validity and reliability of an OSCE in dentistry. *Eur J Dent Educ* 1999;3(3):117-25.
11. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using an objective structured clinical examination. *Br Med J* 1975;1:447-51.
12. Näpänkangas R, Karaharju-Suvanto T, Pyörälä E, et al. Can the results of the OSCE predict the results of clinical assessment in dental education? *Eur J Dent Educ* 2016;20(1):3-8.
13. Cannick GF, Horowitz AM, Garr DR, et al. Use of the OSCE to evaluate brief communication skills training for dental students. *J Dent Educ* 2007;71(9):1203-9.
14. Alaki SM, Shinawi LA, Yamani I, et al. Gathering validity evidence in the use of multiple mini-interviews as an admission tool for dental students: preliminary evidence from Saudi Arabia. *Med Teach* 2016;38(Suppl 1):S45-51.
15. Graham R, Bitzer LAZ, Anderson OR. Reliability and predictive validity of a comprehensive preclinical OSCE in dental education. *J Dent Educ* 2013;77(2):161-7.