A Mini-OSCE for Formative Assessment of Diagnostic and Radiographic Skills at a Dental College in India

Shailesh M. Lele, M.D.S.

Abstract: The objective structured clinical examination (OSCE) is being widely used for assessment of skills in dental education around the world. In India, OSCE awareness is rising, and a few exploratory attempts have been made in its implementation. This article describes use of a five-station mini-OSCE for formative assessment of dental diagnostic and radiographic skills in an undergraduate curriculum. Besides gaining experience in OSCEs, the purpose of this project was to study their validity, objectivity, feasibility, acceptability to students and faculty, and impact on student performance. The mini-OSCE was found to be a fairly valid and reliable tool for formative assessment, though it required more planning, preparation, and resources than other means of assessment. A specially developed orientation module improved its feasibility. The nineteen students perceived it to be a meaningful examination and a fair method due to uniformity of tasks and time allocation; they found the scoring to be transparent and objective. The specific and immediate feedback received was appreciated by both students and faculty members.

Dr. Lele is Principal, Professor, and Department Head, Oral Medicine and Radiology, Sinhgad Dental College and Hospital, India. Direct correspondence and requests for reprints to Dr. Shailesh M. Lele, Sinhgad Dental College and Hospital, S. No. 44/1, Off Sinhgad Road, Pune 411041, Maharashtra, India; 91-20-24351303 phone; 91-20-24351308 fax; lelesgs@vsnl.net.

Keywords: dental education, diagnostic and radiographic skills, assessment, OSCE, India

Submitted for publication 3/4/11; accepted 5/6/11

Observation of students’ performance in the clinical setting is the most frequently used assessment method in health professionals’ courses. According to Oumachigui, this type of assessment has a number of inherent deficiencies in reliability, objectivity, variability, and the limited competencies examined. The objective structured clinical examination (OSCE), first described in 1975 by Harden et al., can overcome these deficiencies. Boursicot and Roberts elaborated on implementation of an OSCE in an undergraduate medical course, and the Dental Council of India in its revised regulations recommends use of the OSCE or objective structured practical examination (OSPE) as a tool for assessing dental students’ skills.

There has been a great deal of interest in this assessment method in dental education worldwide, with many universities and fellowship boards adopting it. In 1994, the National Dental Examining Board of Canada introduced an OSCE into the certification process. At the Royal London School of Medicine and Dentistry, OSCEs have been used for summative in-course assessments, and at Leeds Dental Institute in the United Kingdom they have been used to provide feedback in restorative dentistry. Mossey et al. compared different types of clinical operative skill scenarios in multi-station OSCEs. They found OSCEs were useful in assessment of diagnostic, interpretation, and treatment planning skills, while having limitations in assessment of operative procedures. In 2002, Zartman et al. reported on their seven years of experience with OSCEs in the Department of Pediatric Dentistry at Baylor College of Dentistry. They felt that although time-consuming and labor-intensive, OSCEs provide “unprecedented feedback about students’ understanding and pinpoint areas of confusion.” The OSCE has also been judged to be a valuable method in assessing students’ progress toward competence. Cannick et al. used OSCEs to evaluate dental students’ competence in interpersonal and tobacco cessation communication skills. Derringer surveyed twelve dental schools for assessment methods in the undergraduate course in the United Kingdom, and six of the schools surveyed had adopted OSCEs as the assessment method in orthodontics. In India, Bapat developed and implemented an OSCE in periodontology for final-year B.D.S. students. The majority of students in that study felt that the method was innovative, comprehensive, well structured, objective, and fair and provided scope for self-evaluation; but on the negative side, students found the time allotted to be inadequate and the whole experience to be stressful.

Undergraduate dental education in India is currently going through a transitional period, with the old curriculum changing to satisfy the revised...
regulations released by the Dental Council of India in 2007. The new regulations envision a dental graduate competent to investigate, diagnose, manage, and prevent oral diseases prevalent in India. Two generic skills every dental practitioner must exhibit are 1) clinical evaluation and diagnosis and 2) acquisition of dental radiographic images and their interpretation. Oral Medicine and Radiology is one of the eight clinical subjects in the five-year undergraduate curriculum. At the end of training in this subject, students are expected to demonstrate competence in these two skills.

Training in this subject is introduced in the third year of the course. Besides the didactic component, the third year also sees students getting posted twice to this department. By the end of the second clinical posting, students are expected to be fairly competent in the basic diagnostic process, in making a periapical radiograph (image acquisition and processing), and in its basic interpretation. Formative assessment consists of allotting a case to every student, who then proceeds to record case history; examine the patient; make a provisional diagnosis; advise and make a periapical radiograph; and interpret it. At the end of the allotted time, students present the case findings, diagnosis, radiograph made, and their report. These are evaluated by the faculty, and marks (grades) are awarded. Each of these procedures is made up of several subskills, which the students are expected to perform correctly and in a systematic manner. Mostly, the faculty member does not directly observe the actual performance of these subskills by the students. The very purpose of the formative assessment is thus not fully served.

This kind of global assessment reduces its validity, objectivity, and reliability. Several other drawbacks of this method are that there may be unequal difficulty levels due to lack of uniformity in patient characteristics; all the important skills may not be tested; unequal time may be available to perform the procedures; and at times, patients may exhibit unease and noncooperation. Students too have their grievances about this method, viz. feeling that what they did well went unobserved while other students were getting away with mistakes and perceiving a lack of clarity about what skill is being tested and the level of performance expected.

OSCEs have been shown not only to overcome many of these drawbacks, but also to improve the validity and reliability of assessment. This method has not been tried very often in the undergraduate curriculum in India, particularly so in the area of Oral Medicine and Radiology. Experience and experimentation across various subjects in the dental course would result in refinement of the OSCE. Guidelines on the skills to be assessed and types of scenarios that could be used in OSCEs could then be determined. It was therefore felt that an OSCE could be tried for formative assessment of students’ diagnostic and radiographic skills. Aspects such as its validity, objectivity, feasibility, acceptability to students and faculty, and impact on students’ performance were studied. Faculty training in OSCE implementation was considered to be an added benefit.

Methods

The study was carried out in the Department of Oral Medicine and Radiology at Bharati Vidyapeeth University Dental College and Hospital (Pune, India) between February and May 2010. The project was a curricular requirement for the Certificate Course in Advanced Health Sciences’ Education and Technology conducted by Maharashtra University of Health Sciences (Nashik, India). It was approved by the Institutional Review Board and exempted from ethical review subject to certain conditions. The review board raised concerns about the effect of the OSCE on students’ performance, especially since post-end marks are added to internal assessment marks. The board thus approved the project and agreed to exempt it from ethical review only on ensuring that the process of informed consent would be adhered to, students would be assessed the traditional way and these marks would be added to the internal assessment marks, and marks obtained in the OSCE would be used only for giving feedback to students.

Since faculty members were already sensitized to OSCEs (at a faculty development program held earlier on basic educational science and technology), the author conducted a one-day workshop to review the basics of OSCEs and to develop an implementable scheme. This included specifying domain-specific objectives to be tested based upon the training provided in the posting, identifying appropriate tasks and resources required, and developing checklists and standard setting, instructions to candidates, and overall logistics. Considering that only a few skills were to be assessed and that an OSCE was being implemented for the first time, only five stations were planned (thus, a mini-OSCE).
A trial run carried out with interns was recorded and reviewed for improving the scheme. The recording was also used to develop an orientation module for the students. On the penultimate day of the clinical posting, third-year second-term students were given a short presentation and printed information sheet about the project. After their questions were answered, they gave informed consent to participate in the mini-OSCE. Those who agreed to participate (n=19) were presented with the Orientation Module consisting of text slides, pictures of the setting, and video clips of the scheme.

The mini-OSCE consisted of five stations:

- Station 1 (Procedure): communication skill. Task: recording of a specified component of patient history. (5 minutes, 10 marks)
- Station 2 (Procedure): patient examination skills. Task: clinical examination of specified dental tissues or structures. (10 minutes, 20 marks)
- Station 3 (Response): answering case scenario-based questions. (5 minutes, 10 marks)
- Station 4 (Response): interpretation of a periapical radiograph. (10 minutes, 15 marks)
- Station 5 (Procedure): radiography technique. Task: demonstrate bisecting the angle technique of intraoral periapical radiography of specified tooth. (6 minutes, 15 marks)

Stations 1 to 4 were set up in the Undergraduate Clinic. Five-minute holding stations were placed between Stations 1 and 2 and between Stations 3 and 4 to synchronize student movement. Station 5 was set separately in the Radiology Clinic. Eight students were taken through at a time. While four rotated through the stations in the Undergraduate Clinic, the other four performed in the Radiology Clinic. After one cycle, students were interchanged. Rotation through Stations 1 to 4 required forty minutes, while the total time required at Station 5 was about thirty minutes. Adding the preparation and clean-up time, one cycle could be completed in about one hour. Thus, in two hours’ duration eight students could be assessed in two cycles.

One faculty member was positioned at each of the Stations 1, 2, and 5 who assessed students’ performance using a checklist. Stations 3 and 4 did not require observers, but students’ responses at these two stations were assessed using a checklist. In addition to four faculty members, the services of three postgraduate students, one staff nurse, one radiography technician, and one attendant were utilized for tasks such as station preparation and clean-up, selection of patients/volunteers, recording of patient features, and time-keeping.

Immediately after the examination, the results were compiled and displayed on the board, and a structured, stationwise feedback was provided to students individually. After feedback was given to the students, their feedback was obtained with a structured questionnaire. Later, on completion of mini-OSCE implementation, feedback was obtained from the faculty with another structured questionnaire.

### Results and Discussion

The study was planned for four groups of third-year second-term students, for a total of forty. However, only nineteen participated: one student was absent, one group of ten declined to participate, and another group of ten was not available due to unexpected change in posting schedule. These nineteen students’ performance data (Table 1) and feedback were analyzed. Feedback was also obtained from four faculty members who participated.

The overall mean score (52.1 percent) on the mini-OSCE was just over the passing standard (50 percent). This was due to students’ poor performance at Stations 2, 3, and 4 (mean scores: 40.5 percent, 43.9 percent, 45.1 percent), which could be attributed to lack of sufficient knowledge and experience in identification of a disease and interpretation of clinical or radiographic data. It could also be due to deficiencies in the demonstration of these procedures by the faculty. The poor performance at these stations

---

**Table 1. Students' performance on the mini-OSCE, by percentage correct at each of the five stations and total for all stations**

<table>
<thead>
<tr>
<th>Mean Marks Obtained at Station</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>77.4%</td>
<td>40.5%</td>
<td>43.9%</td>
<td>45.1%</td>
<td>62.8%</td>
<td>52.1%</td>
</tr>
<tr>
<td>±SD</td>
<td>9.2%</td>
<td>13.8%</td>
<td>25.4%</td>
<td>15.2%</td>
<td>18.2%</td>
<td>9.1%</td>
</tr>
</tbody>
</table>
notwithstanding, perusal of checklists showed that students adhered well to the procedures and a systematic approach. The good mean score (77.4 percent) at Station 1 could be attributed to the fact that it was a relatively easy task that was less dependent than the others on experience, while the good score (62.8 percent) at Station 5 could be due to relatively more experience in the task during the posting. Also, most of the students were well aware by this time of the skills required in practice—radiography being one of them—and they were more interested in learning such skills.

Validity

Validity is a unitary concept requiring evidence from a variety of sources (content, response process, internal structure, relationship to other variables, and consequences). Table 2 shows how far the “questions” (the station tasks) matched the learning objective being tested. In this particular posting, students are trained in arriving at a provisional diagnosis based on the case history and patient examination data and also in making a periapical radiograph and interpreting it. These are the two most important diagnostic skills required in dental practice. The first skill requires the student to achieve objectives under affective, cognitive, and psychomotor/communication domains. Unlike the medical practitioner, the dental practitioner is required to actually make radiographs at various stages of diagnosis, treatment, and follow-up and to interpret them. Therefore, this skill assumes almost equal importance as the first skill. This skill also requires the student to achieve objectives under affective, cognitive, and psychomotor/communication domains. The checklists were prepared to assess objectives under all three domains.

At face value, the weight allocated to the broad objectives and the specific objectives appears to match that in real practice. In responses to the evaluation questionnaire, all the faculty members felt that “the objectives chosen to be assessed were appropriate to the skills learnt during the posting” and that “the task to be performed at each station was appropriate for the objective being assessed.” Eighty-nine percent and 79 percent of the students, respectively, agreed with these two statements. The items in the checklists were judged by the faculty to be appropriate. While all the faculty members agreed that the time allotted was adequate at the five stations, 68 percent of the students felt it was inadequate at Station 4. All the students felt that the Orientation Module made them familiar with the new format, which was reflected in smooth conduct of the examination.

These students were also assessed the traditional way in the radiographic technique exercise. At Station 5 in the mini-OSCE, students were asked to demonstrate the technique without actually exposing the film, whereas in the traditional examination students were required to not only expose but also process the film. The resultant radiograph was then assessed and marks given. Table 3 shows the scores obtained by the students in these two examinations.

Even though the mean scores in the two assessment formats were similar, the range of scores was much wider in the mini-OSCE. The assessment by the traditional method tends to be global with only the outcome rather than the process being assessed. It may happen that steps missed or incorrectly performed by a student get overlooked or forgiven by a generous examiner or some steps correctly performed may not get assessed. In the OSCE, on the other hand, all the steps in the process are observed and assessed. The OSCE, thus, could be said to measure what it intends to measure. The wider range of scores in the OSCE reflected its ability to discriminate among poor, average, and high performers. Due to the global

<table>
<thead>
<tr>
<th>Topics</th>
<th>LO</th>
<th>Stationwise Marks</th>
<th>Marks per LO</th>
<th>% per LO</th>
<th>% per Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arriving at a provisional diagnosis</td>
<td>RCH</td>
<td>10 - - - - -</td>
<td>10</td>
<td>14.29%</td>
<td>57.14%</td>
</tr>
<tr>
<td></td>
<td>PE</td>
<td>– 20 - - - -</td>
<td>20</td>
<td>28.57%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICD</td>
<td>– - 10 - -</td>
<td>10</td>
<td>14.29%</td>
<td></td>
</tr>
<tr>
<td>Periapical radiology</td>
<td>IOPA-T</td>
<td>– - - - 15</td>
<td>15</td>
<td>21.43%</td>
<td>42.86%</td>
</tr>
<tr>
<td></td>
<td>IR</td>
<td>– - 15 - -</td>
<td>15</td>
<td>21.43%</td>
<td></td>
</tr>
</tbody>
</table>

LO=learning objective, RCH=recording case history, PE=patient examination, ICD=interpretation of clinical data, IOPA-T=introral periapical radiograph technique, IR=interpretation of radiograph.
assessment, scores in the traditional method of assessment tend to be closer to the mean.

Use of a checklist in this OSCE enabled the faculty members to provide student-specific and procedure-specific feedback. This was much appreciated by the students, with 89 percent of them able to identify their weak areas. Twelve (63 percent) students felt motivated to learn further after OSCE experience.

Reliability

According to Downing, reproducibility of assessment data or scores over time or occasions is a major source of validity evidence. In the split-half methodology of reliability estimation, items in a test are split into two tests that are equivalent in content and difficulty. If each station could be considered as a test, the items in the checklist could be analyzed by calculating the Pearson product moment correlation coefficient, which indicates internal consistency. The reliability of the whole test can then be predicted using the Spearman-Brown formula. Table 4 shows these statistics.

For Stations 1 and 2, reliability was calculated separately for two groups since the items in the checklist were qualitatively different. For Station 5, identical checklists were used for the two groups. The reliability of these three stations could be considered adequate, especially since this assessment was formative in nature. For such assessments with lower consequences, created and administered by faculty members, Downing expects reliability in the “range of 0.70–0.79 or so.”

Feasibility

The mini-OSCE was successfully implemented, with considerable input from the faculty, a lot of brainstorming, and much planning involved. Its feasibility depended on the active involvement of faculty members at every stage, along with the supporting staff. No additional infrastructure and time for actual implementation, however, were required. The one-day workshop and the trial run with interns were found by the faculty adequate to help them understand the concepts and implementation of an OSCE. Similarly, all the students felt that the Orientation Module was adequate to help them understand this new form of examination. Both these responses increase the feasibility of conducting an OSCE. However, selection of patients and volunteers was found to be a critical factor. Also, at every stage of the examination, the human resource requirement was twice that for the traditional method.

Acceptability to Students and Faculty

If a new form of examination overcomes the problems that the students have with the traditional examination, it could be considered acceptable to them. Analysis of students’ feedback on this OSCE revealed the following: all the students felt that the scoring was transparent and objective; 95 percent felt that it was well organized, well structured, and well sequenced; 89 percent felt that it was meaningful since important skills learnt during the posting were assessed; 89 percent of them

<table>
<thead>
<tr>
<th>Test</th>
<th>Score (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (±SD)</td>
<td>Minimum</td>
<td>Maximum</td>
<td>Range</td>
</tr>
<tr>
<td>Traditional</td>
<td>66.1% (±12.3%)</td>
<td>50.0%</td>
<td>85.0%</td>
<td>35.0%</td>
</tr>
<tr>
<td>Mini-OSCE</td>
<td>62.8% (±18.2%)</td>
<td>30.0%</td>
<td>90.0%</td>
<td>60.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calculation</th>
<th>Station 1 1st Group</th>
<th>Station 1 2nd Group</th>
<th>Station 2 1st Group</th>
<th>Station 2 2nd Group</th>
<th>Station 5 Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson product moment correlation coefficient</td>
<td>0.53</td>
<td>0.75</td>
<td>0.60</td>
<td>0.48</td>
<td>0.70</td>
</tr>
<tr>
<td>Spearman-Brown formula</td>
<td>0.69</td>
<td>0.86</td>
<td>0.75</td>
<td>0.65</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Table 3. Comparison of scores obtained in the radiographic technique exercise

Table 4. Reliability estimation for Stations 1, 2, and 5
were able to identify their weak areas; and 63 percent felt motivated to learn further after OSCE experience.

While this feedback shows a strong possibility that students will accept the OSCE, it should also be noted that 63 percent of the students felt that this format was more stressful than the traditional format and 79 percent felt that they were frightened/scared when performing in a faculty member’s presence. All the educators had similar observations to make about their perceptions of the students. It would be interesting to research if the performance would differ if it was videotaped (as suggested by some students and faculty members) rather than directly observed. At the Department of Pediatric Dentistry at Baylor College of Dentistry, the students’ anxiety was overcome by holding focus group sessions and giving them practice OSCE modules. The comments provided in the feedback on my OSCE give the impression that students would prefer to experience an OSCE during the posting rather than at the end because of the feedback it provides.

All or the majority of the faculty members involved in implementing this OSCE felt that the OSCE can assess a wide range of objectives under the three domains; use of a checklist was helpful in being more objective and may reduce interexaminer variation; the OSCE makes it possible to provide immediate and specific feedback to students; and it was not tiring or boring to be an examiner. However, the majority also felt that the OSCE is appropriate only for posting-end assessment and not for summative assessment.

Impact on Students’ Performance

Thirty-seven students in the same class had previously undergone the traditional form of assessment in which eliciting and recording the case history was the exercise, and their scores were available for comparison with the students who participated in the OSCE. Assuming a similar mix of students as far as this skill is concerned, it could be seen that students performed much better on the OSCE (Table 5). Not only was the score range narrower on the OSCE, but there was a shift of scores to higher values.

It appears that the OSCE does not necessarily improve student scores (Table 3). However, the wide range of scores probably indicates that more critical assessment that occurs with the checklist punishes the poor performers while rewarding the high performers. It is unclear whether these two interpretations need to be reconciled or could be treated independently.

Conclusions

The mini-OSCE was found in this study to be a fairly valid, reliable, and feasible assessment method, particularly for formative assessment. However, a larger sample size would perhaps have provided results more representative of the population under study. Using a checklist made the assessment much easier, but calculation of marks later needs to be done very meticulously. The OSCE provided specific and immediate feedback to students on their performance and to teachers on their educational methods, making timely correctional measures possible. Prior training in the OSCE eased the task of convincing faculty members to adopt this method of assessment. Selecting patients/volunteers with characteristics suitable for the tasks at procedural stations is a daunting task. If a station is to be visited by more than four students, a larger number of suitable patients/volunteers is required to be on hand. Lack of interaction (chairside viva) with the students during the examination is not liked by examiners. Students can and do decline to participate in something new, but this should be accepted in the right spirit. Conducting educational research is very different from clinical trials. Much more conceptualization and planning are required for an OSCE than for other assessment methods, so a great deal of support is needed from one’s colleagues to carry out a study of this nature.

Acknowledgments

The author thanks Dr. Payal Bansal (In-Charge, MET Department, Maharashtra University of Health

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean (±SD)</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>61.2% (±9.8%)</td>
<td>40.0%</td>
<td>80.0%</td>
<td>40.0%</td>
</tr>
<tr>
<td>Mini-OSCE</td>
<td>77.4% (±9.2%)</td>
<td>65.0%</td>
<td>95.0%</td>
<td>30.0%</td>
</tr>
</tbody>
</table>
Sciences) for mentorship throughout the project; the faculty of the Department of Oral Medicine and Radiology at BVU Dental College and Hospital for their cooperation; and the students who consented to participate in the project, all at Pune, India.

REFERENCES