

The Comparison of OSPE with Conventional Physiology Practical Assessment

Aarti Sood Mahajan, M.B.B.S, M.D.,¹ Nilima Shankar, M.B.B.S, M.D.² and
O.P.Tandon, M.B.B.S, M.D, M.N.A.M.S, F.A.M.S.²

¹Department of Physiology, Maulana Azad Medical College,
Bahadur Shah Zafar Marg, New Delhi 110002 INDIA

²Department of Physiology, University College of Medical Sciences and G.T.B. Hospital
Dilshad Garden, Delhi 110095 INDIA

Phone: (+)91-011-6123784

Fax: (+)91-011-6132011

Email: aartis_mahajan@yahoo.co.in

ABSTRACT

The Objective Structured Practical Examination (OSPE) is a new concept in practical assessment of physiology in our country. It is a modified form of Objective Structured Clinical Examination (OSCE) but is used for evaluation of pre and paraclinical subjects. Although theoretically known, very few medical colleges have incorporated OSPE as an assessment tool in the curriculum of first year medical students. We wanted to compare the marks obtained in OSPE and the other conventional methods. The OSPE marks showed similarity with clinical examination and were different from marks of other experimental procedures like graph and chart. A similar result was found in two examinations. We conclude that OSPE can replace the existing pattern of clinical examination. To replace others it would require an elaborate and structured OSPE bank. Presently, it can supplement but not replace the conventional methods. Any change must first be thoroughly evaluated before it can uproot a well-defined and time-tested assessment methodology.

INTRODUCTION

A uniform and reliable practical evaluation of medical students is always desirable. In order to ensure objectivity, we have converted to MCQ's for selecting new entrants to the medical colleges and short structured questions for evaluating theoretical knowledge during the medical curriculum. In line with the same principle and to make practical assessment more comprehensive, objective and unbiased we decided to introduce the Objective Structured Practical Examination (OSPE) in the first year of our medical college. The OSPE is a method of objectively testing the practical knowledge and skills acquired during the preclinical years of a medical curriculum. This method was standardized by the All India Institute of Medical Sciences.¹ It is a modification of Objective Structured Clinical Examination (OSCE) used for clinical evaluation.² At present OSPE is conducted in a few medical colleges in association with other conventional methods, and is being allotted a small percentage of marks. However, in time to come, it is expected to replace the other subjective assessment methods. Introducing a new concept in a traditional framework is always met with scepticism.³ The aim of the present study was to see the relationship between the scores obtained by medical students in OSPE and other traditional methods during various evaluations in the first

year of medical college. This pilot study was meant to give us an idea whether any of the existing evaluation methods were similar to OSPE and could be replaced in order to increase the objectivity in assessment.

MATERIALS AND METHODS

The study was conducted in the department of physiology on 100 medical students at the University College of Medical Sciences, Delhi, India. Seventy-six students participated in all evaluations held at different times during the first year of the medical curriculum.

The classroom exercise

The students were tested in groups of 25 each. The OSPE consisted of two procedures and two question stations. The questions were changed randomly for different groups of students. An attempt was made to make the atmosphere as congenial as a routine classroom exercise with the aim to familiarize the students with this new system of examination.

The examination exercise

Two examinations were held in different semesters with different course content. The OSPE was included along with the conventional methods of practical assessment. A

Table 1. Comparison between different assessment methods in first examination.

	1B	1C	1D
1A	0.000*	0.000*	0.059
1B		0.028*	0.000*
1C			0.000*

The p value of the difference between the observed means of various groups. * $p < 0.05$ is significant. 1, represent the first exam. A, B, C and D denote marks of OSPE, hematology, graph and chart, clinical examination respectively.

similar set of eight question and two procedure stations were allocated to 25 students. A student spent three minutes at each station, questions were randomly changed each day and the entire schedule lasted for four consecutive days. The OSPE questions were designed to test the cognitive aspect of learning mainly knowledge, comprehension and synthesis of facts. The procedure stations were from clinical examinations and hematology course chosen to evaluate the psychomotor skills and affective domain. Each question was to evaluate a specific learning objective of the course content.

Other conventional assessment techniques included hematology practical (HP), clinical examination (CE) and graphs and charts (GC). In HP the students were asked to perform a small laboratory exercise using a blood sample, like counting blood cells or staining smears. This is not necessarily done in front of the examiner and was followed by an oral question answer session (viva voce). In CE the students were asked to clinically examine and elicit signs in a simulated patient. The idea was to test the techniques of doing a clinical examination of various systems in simulated patients. The examiner evaluated the psychomotor skills and asked relevant questions. In GCs, some diagrams, photographs and graphs were given to the student to explain. All these exercises were followed by a viva or a question answer session. A total of ten marks were allocated to each type of assessment procedure. All these evaluation types have been routinely practiced for many years in most medical colleges of our country. They are relatively subjective, unstructured, can have errors of bias, ambiguity and obsolesce. It is for this reason that the concept of OSPE was introduced.

Statistical Analysis

The mean and standard deviation of marks obtained in all exercises (hereafter referred as groups) were calculated. Hierarchical analysis of variance showed the change between the groups was significant. ($f=48.842$, $p < 0.001$). Tukey's test of multiple comparisons at 5% interval based on observed means was done to individually compare the result of the different methods of assessment. (table 1 and 2).

Table 2. Comparison between different assessment methods in second examination.

	2B	2C	2D
2A	0.105	0.000*	0.374
2B		0.442	0.000*
2C			0.000*

The p value of the difference between the observed means of various groups. * $p < 0.05$ is significant. 2 represents the second exam. A, B, C and D denote marks of OSPE, hematology, graph and chart, clinical examination respectively.

Comparison was done between any two groups at 95% confidence level ($p < 0.05$ is significant).

Simple Pearson's correlation method was used to find the association between marks of the question and procedure stations. The mean difference between the two was compared by the Student's t test.

RESULTS

The mean and standard deviation of marks obtained in all exercises is shown in Figure 1. In the second examination (4-8) the scores obtained are higher.

Comparison of OSPE with hematology practical

The OSPE marks of first examination were different, but those of the second exam were similar to the respective HP marks. (Table 1 and 2).

Comparison of OSPE with graphs and charts

The OSPE and corresponding GC marks of both examinations were statistically different ($p < 0.05$). (Table 1 and 2).

Comparison of OSPE with clinical examination.

The OSPE and corresponding CE marks of both examinations were similar. (Table 1 and 2).

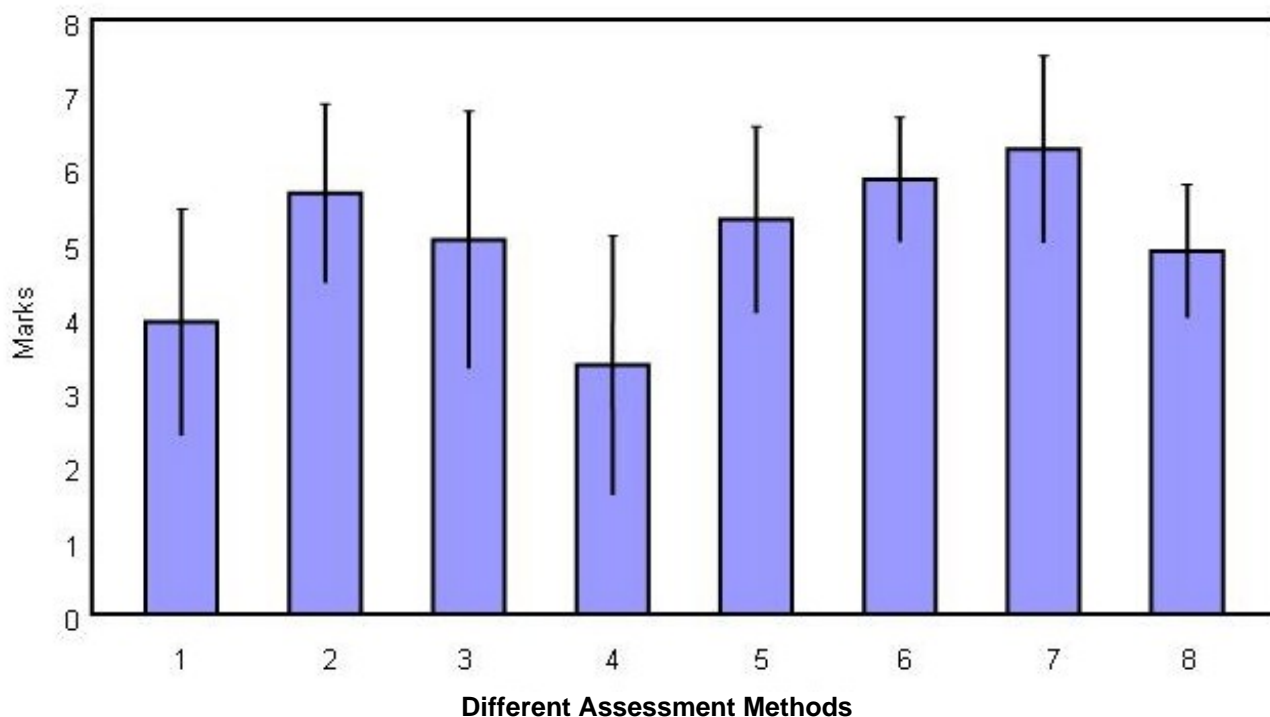
Comparison between other groups

Similarity was seen between marks of HP and GC of second exam. All other comparisons showed differences.

Comparison between question station and procedure station of OSPE of classroom exercise.

No significant correlation was observed between the two question stations, two procedure stations, and the total marks obtained in the question and procedure stations. ($p > 0.05$). When the mean marks were compared by the two-tailed 't' test there was significant difference between the two questions and procedure stations ($p = 0.008$, $p = 0.001$ respectively), but the total marks of both the question and procedure stations did not differ significantly ($p > 0.05$).

Figure 1. Marks of Different Assessment Methods



The mean and standard deviation (STD-Dev) of the marks obtained in different forms of assessment. Maximum marks =10. 1, 2, 3,4 represent marks of OSPE, hematology, graph and chart, clinical examination respectively in the first examination.5, 6,7,8 represent marks of OSPE, hematology, graph and chart, clinical examination in the second exam.

Students scored differently in the various evaluation procedures. OSPE marks were similar to clinical examination and different from marks of graphs and charts. The comparison with other forms of evaluation is varied.

DISCUSSION

In an attempt to improve the practical assessment in our institution, OSPE was introduced for the first time along with the other conventional assessment procedures like HP, GC and CE in the first year of medical curriculum. We wanted to see the relationship of marks obtained in OSPE and other forms of evaluation.

Our first observation was that a significant variation in the marks obtained in different assessment procedures. This was a consistent finding and was found in both the examinations. We generally expect that good students would do well in any form of evaluation.⁴ Since this was not observed we presume that the various assessment methods assess different capabilities of the students.⁵ The comparison of OSPE with other forms of assessment, showed a consistent similarity to CE, a variation with GC and a varied response to HP. In procedure stations of OSPE and CE the student performs in front of the examiner and the psychomotor skills are mainly tested. In a previous study, undertaken to see the relationship between OSCE and clinical cases, no correlation was reported. It was suggested that OSCE should be

employed for evaluation of specific clinical skills, but for comprehensive evaluation a combination of OSCE and clinical cases should be used.⁶ In the initial part of our checklist we had included statements to test the affective domain of learning. The student had to address the subject politely, make him comfortable, explain the procedure etc. The question stations were included to evaluate the cognitive aspect of learning. A written component, when added to OSCE, is known to improve reliability and economize on the resources.⁷ Our question stations in OSPE also served this purpose. We are of the opinion that OSPE if properly structured, along with a short written component can replace the current clinical examination exercise taught in the preclinical years. It may not be so useful in the final year course if the investigation, differential diagnosis and management are to be discussed unless different OSPE (OSCE) stations are made for each.

OSPE scores were different from those of GC. These test the student in cognitive aspects like recall and interpretation. They also rely on the communication skills of the student. Question stations would have to be elaborately designed if they are to be used instead of GC. When other assessment methods were compared with each other, a varied pattern of responses was observed. The lack of correlation between both the procedure and the question stations could be because they were testing different things. A student may not know everything.

A few questions emerge from this. Firstly, should we expect the different instruments of assessment to yield similar results? Secondly, if they do, does it mean that they are testing the same domain and in that case, should they be continued together or not?

To answer these questions we must look into the association of OSCE and other assessment tools. OSCE has shown positive correlation with other forms of assessment like ward evaluation, American Board of Surgery In training Examination (ABSITE), short answer questions and subjective rating.^{8,9} However basic science scores and MCQ's showed no correlation to OSCE.^{10,11} Hence there is no generalization that all assessment tools must correlate. If there is similarity between the instruments it could be that we are testing the same skills.¹² A similar study has been carried out earlier.⁵ The authors found no correlation between varied forms of assessment and felt that these methods were testing different abilities of the students. We must however keep in mind that there is no gold standard for assessment and so we cannot say which method is better.⁸ The criterion of a good examination includes validity, reliability, objectivity, practicability, relevance, and promotion of learning, power to discriminate between students, relaxed environment and a positive student feedback.¹³ Clearly no single test fulfills the criterion of a good examination and the different methods complement each other.

If we are to discontinue any method we must be aware of the possible consequences on learning.⁸ The OSPE-like OSCE is associated with "achieving" style of learning but not with "meaning" or "reproducing" style of learning.¹⁴ The OSPE complements other methods of assessment. It allows us to directly observe the student, give similar questions to all students, check on minute details in order to standardize and focus our evaluation, to be more objective and unbiased in marking. On the other hand our conventional methods allow for an in-depth analysis of the subject, with more interaction between the examiner and the student. The examiner's professional judgment and experience can make the examination a learning exercise as it provides an instant feedback to the student. These advantages justify their inclusion.

CONCLUSIONS

Our results have shown OSPE marks similar to clinical examination and different from graphs and charts. In the present set up OSPE can replace clinical examination and not graph and charts. However we feel that an elaborate OSPE bank may be able to overcome these lacunae. Till that is done, in physiology practical examination we should utilize different techniques in order to increase the validity of the examination.

REFERENCES

1. Bijlani, R.L. and Nayar, U. Objective structured practical examination. *Teaching Physiology, Trends and Tools*. 1983; 151-159.
2. Harden, R.M and Gleeson, F.A. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Medical Education*. 1979; 13: 41-54.
3. Kowlowitz, V, Hoole, A.J, and Sloane, P.D. Implementing the objective structured clinical examination in a traditional medical school. *Academic Medicine*. 1991; 66(6): 345-347.
4. Newble, D.I and Swanson, D.B. Psychometric characteristics of the objective structured clinical examination. *Medical Education*. 1988; 22: 325-334.
5. Nayar, U, Malik, S.L, and Bijlani, R.L. Objective structured practical examination: a new concept in assessment of laboratory exercise in preclinical sciences. *Medical Education*. 1986; 20: 204-209.
6. Verma, M and Singh, T. Experience with the Objective structured clinical examination (OSCE) as a tool for formative evaluation in pediatrics. *Indian Pediatrics*. 1993; 30: 699-702.
7. Verhoeven, B.H, Hamers, J.G.H.C, Scherpbier, A.J.J.A, Hoogenboom, R.J.I and Van der Vleuten C.P.M. The effect on reliability of adding a separate written assessment component to an Objective structured clinical examination. *Medical Education*. 2000; 34: 525-529.
8. Collins, J.P and Gamble, G.D. A multiformat interdisciplinary final examination. *Medical Education*. 1996; 30: 259-265.
9. Schwartz, R.W, Donnelly, M.B, Sloan, D.A, Johnson, S.B and Strodel, W.E. The relationship between faculty ward evaluation, OSCE, and ABSITE as measures of surgical intern performance. *The American Journal of surgery*. 1995; 169: 414-417.
10. Sloan, D.A, Donnelly, M.B, Schwartz, R.W, and Strodel, W.E. The Objective structured clinical examination. *Annals of Surgery*. 1995; 222(6): 735-742.
11. Johnson, G and Reynard, K. Assessment of an Objective structured clinical examination (OSCE) for undergraduate students in accident and emergency medicine. *Journal of Accident and Emergency medicine*. 1994; 11: 223-226.
12. Roberts, J and Norman, G. Reliability and learning from an objective structured clinical examination. *Medical Education*. 1990; 24: 219-223.
13. Paul, V.K. Assessment of clinical competence of undergraduate medical students. *Indian Journal of Pediatrics*. 1994; 61: 145-151.
14. Martin, I.G, Stark, P and Jolly, B. Benefiting from clinical experience: The influence of learning style and clinical experience on performance in an undergraduate objective structured clinical examination. *Medical Education*. 2000; 34: 530-534.