

# Traditional clinical examination vs objective structured practical examination in human physiology: Examiner's bias

Pinaki D Wani

Department of Physiology, KJ Somaiya Medical College and Research Centre, Sion, Mumbai, Maharashtra, India.  
Correspondence to: Pinaki D Wani, E-mail: wanipinaki@gmail.com

Received April 17, 2014. Accepted February 11, 2015.

## Abstract

**Background:** A traditional clinical examination (TCE) in physiology focuses on the “knows” and “knows how” aspects and is inadequate in evaluating the overall performance of the students. Objective structured practical examination (OSPE) focuses on the “shows how” aspect of Miller’s pyramid of competence.

**Objective:** This study was aimed to compare and establish the relation of the marks given by examiners with varied teaching experiences in determining the validity and reliability of TCE and OSPE.

**Materials and Methods:** A total of 50 first-year MBBS students in Physiology were divided in two batches of 25 each. These were further divided in three batches with eight, eight, and nine students examined by three examiners with teaching experience of 35 years, 6 years, and 1 year, respectively, in each batch. All the examiners conducted TCE followed by OSPE for the same batch of students for four modules in abdominal system.

**Results:** Analysis of variance and *post hoc* Bonferroni’s test for comparative study among examiners in TCE showed significant difference (module A,  $P = 0.000$ ; module B,  $P = 0.000$ ) in the mean marks given by all the three examiners while OSPE showed no significant difference (module A,  $P = 0.829$ ; module B,  $P = 0.842$ ) in the mean marks given by all the examiners for all the modules.


**Conclusion:** OSPE is an objective, structured, unbiased assessment method that can be incorporated with less experienced examiners along with the TCE for overall assessment of the students.

**KEY WORDS:** Objective structured practical examination (OSPE); traditional clinical examination (TCE); reliability; validity; examiner’s bias

## Introduction

The Graduate Medical Education curriculum in the South-East Asia region is moving its focus from knowledge to competency-based education, hence the assessment

of students’ clinical competencies becomes an integral part of the medical curriculum.<sup>[1]</sup> Assessment can be defined as the process of drawing inferences from students’ work and estimating his/her worth in terms of marks.<sup>[2]</sup> The format of assessment influences the learning processes in the students. A modification in assessment method can improve the quality of medical education.<sup>[3,4]</sup> Newble and Jaeger<sup>[5]</sup> described that by changing clinical assessment in the final year from a pass/fail system based on ward reports to a clinical practical examination increased the time spent by medical students in the wards. A good assessment tool must fulfill criteria’s of objectivity, validity, reliability, and feasibility. A reliable examination must be valid, objective, and unbiased to be acceptable.<sup>[6]</sup> A traditional clinical examination (TCE) in Physiology involves performing a particular clinical

Access this article online	
Website: <a href="http://www.ijmsph.com">http://www.ijmsph.com</a>	Quick Response Code: 
DOI: 10.5455/ijmsph.2015.17042014122	

International Journal of Medical Science and Public Health Online 2015. © 2015 Pinaki D Wani. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for any purpose, even commercially, provided the original work is properly cited and states its license.

procedure and a bedside viva voce, which is followed by the assessment-based on global performance rather than candidate's individual clinical competency. It mainly focuses on the "knows" and "knows how" aspects of Miller's pyramid of competence.<sup>[7,8]</sup> This can often be subjective, biased, monotonous, and inadequate in evaluating the overall performance of the students at all levels of knowledge, skill, and attitude.<sup>[9]</sup> Often attitude, communication skills, interpersonal skills, ethical issues, and professional judgments are not tested in TCE and the final conclusions made by the students are questioned rather than emphasizing on the procedure and clinical examination.<sup>[10-12]</sup> In addition, the traditional assessment tools test facts and recall knowledge but fail to assess the understanding of core topic and the problem-solving skills of the students.<sup>[12]</sup> Students are assessed by examiners with varied teaching experiences, and this increases examiner's subjectivity and reduces the reliability of the examination. This examiner's subjectivity reduces the correlation coefficient between marks given to the same candidate's performance to as low as 0.25, which significantly affects the scoring, resulting in dissatisfaction among both the examiners and examinees.<sup>[7]</sup> In addition, TCE method lacks a proper feedback process to the students in order to improve their skills.<sup>[12]</sup> In theory examinations, these issues are effectively tackled in the form of incorporating multiple choice questions (MCQs) and short answer questions. Similar requirements in practical examinations resulted in introducing OSCE/OSPE as a more objective, structured, and unbiased assessment tool that is reliable and valid. Objective structured practical examination (OSPE) was derived from objective structured clinical examination (OSCE) and modified by Harden and Gleeson (1975-79).<sup>[7]</sup> The OSPE assesses practical competencies in a methodical, objective, and time-orientated manner with direct observation of student's performance during planned clinical test stations.<sup>[13-15]</sup> It assesses the third "shows how" level of Miller's pyramid focusing on assessment of performance of specific skills in a controlled setting, which makes it relevant as an assessment tool in undergraduate years of learning.<sup>[8]</sup> The standardized checklists ensure maximum objectivity and reliability in an OSPE.<sup>[1,16]</sup> However, in today's time, OSPE is conducted in limited medical universities all over India and is allotted less percentage of the marks.<sup>[17]</sup> Also, there have been few studies concentrating on the role of examiners with different teaching experiences on the reliability and validity of OSPE as an assessment tool. Hence, the present study was aimed to compare and establish the relation between the scores given by the examiners with varied teaching experiences in TCE and OSPE. The aim of this study was to understand the role of examiners with different teaching experiences in determining the validity of TCE and OSPE.

## Materials and Methods

Human Physiology is taught in the first two terms of MBBS curriculum under MCI (Medical Council of India)

regulations. The study was conducted on 50 first-year MBBS medical students in the Department of Physiology at the K.J. Somaiya Medical College and Research Centre, Mumbai, India, after obtaining approval from Institutional Ethics Committee for Research on Human Subjects and a written informed consent from the participants. These medical students (14 men and 36 women) with the mean age of  $18.6 \pm 2.4$  years in their second term in Physiology had  $\geq 65\%$  marks in state common entrance test for admission in the medical university. They were introduced to the system of OSPE by a short lecture, power-point presentation and a role-play organized by the faculty members. A full-day workshop was arranged by the Medical Education Unit of K.J. Somaiya Medical College and Research Centre for the faculty members to acquaint them with the newer assessment tools. The workshop had a 3-h interactive session on OSCE/OSPE with demonstration on the formation of OSCE/OSPE stations and a practice session. Following this, blueprint of the structured checklist for observed stations, unobserved stations for four separate modules in abdominal system examination was prepared along with examiner's and student's instruction manual. The questions for OSPE modules were selected as per "must know," "desirable to know," and "nice to know" criteria, and the difficulty level was maintained equal for all the four modules. These were validated by the senior faculty members who had an extensive experience in the teaching field and were active in medical curriculum designing. However, in TCE, no such set pattern was followed. A total of 50 first-year MBBS students were divided in two batches of 25 each. Each batch of 25 students was further divided in three batches with 8, 8, and 9 students in each batch to be examined by three examiners with teaching experience of 35 years, 6 years, and 1 year, respectively, as per their roll numbers. The examiners selected for the study had a basic training in the medical education innovations. All the three examiners conducted TCE followed by OSPE for same batch of students for four separate modules on 4 consecutive practical days in Physiology. On each of the 4 days, each batch of 25 students appeared for TCE of module 1 in abdominal system clinical examination followed by OSPE of module 1 in abdominal system and subsequently similar assessment was performed with modules 2, 3, 4 of TCE and OSPE. This was done to avoid any overlap and bias in performance by the students. Modules 1 + 2 were clubbed as module A and modules 3 + 4 were clubbed as module B for statistical analysis.

Trained standardized participants were allotted for both the assessment tools. In the traditional assessment method, each student performed a clinical skill, which was followed by bedside *viva voce* on the same, and the assessment of each student was done based on the overall performance of the student.

For TCE, all the three examiners were briefed about the content and the flow of the examination but no structured format was made available for their reference. Questions

	Day 1: Roll No 1–25	Day 2: Roll No 26–50	Day 3: Roll No 1–25	Day 4: Roll No 26–50
TCE	Module 1	Module 2	Module 3	Module 4
OSPE	Module 1	Module 2	Module 3	Module 4

**Table 1:** Comparative study among examiners with different teaching experiences

Sources of variation (between groups)	Method	Significance (two-way ANOVA)
Module A (1 + 2)	TCE	0.000*
	OSPE	0.829
Module B (3 + 4)	TCE	0.000*
	OSPE	0.842

\* $P < 0.001$ .

were asked randomly based on the examiner's experience, thoughts, perceptions, and mood. The entire TCE session ended in 90 min.

While with the OSPE, students were oriented by providing an OSPE map and a written instruction list before the start of the examinations, and they moved from one station to another following the audible ring by the timekeeper. In all the four modules of abdominal system, OSPE consisted of 10 stations of 3–5 min including:

1. One station on communication skills (1 mark);
2. 2–3 observation/procedure station on inspection, palpation, and percussion of abdominal system (2–3 marks); and
3. Six unobserved stations with questions related to the procedural stations and MCQs (1–2 marks) + one rest station arranged in Physiology practical laboratory in a clockwise manner.

The entire session lasted for 50–60 min on all 4 days.

Passing cutoff at 50% was decided for both TCE and OSPE based on the criterion referencing and MCI norms.<sup>[17]</sup> An examiner appointed at procedural stations was provided with a prevalidated checklist to mark immediately according to the observed procedure. Statistical analysis of data was performed in the following

way: data were compiled using Microsoft Office 2007 Excel Software and analyzed using SPSS software (Version 16.0).

Examples of unobserved stations in OSPE:

- What is ascites? What are the causes of ascites? (2 ½ mks)
- Why should the patient take deep breaths during palpation of abdomen? (1 mk)

## Results

All the 50 students participated in the study.

## Discussion

In an attempt to modernize our assessment system and make it more competence based, we compared TCE with OSPE in the current study. Despite an extensive work in OSPE, very few studies are conducted to assess examiner's variability in TCE and OSPE in terms of their teaching experience. In the present study, the results of two-way analysis of variance (ANOVA) and *post hoc* Bonferroni's test for comparative study among examiners showed significant difference in the mean marks given by examiners 1 (35-year teaching experience) and 3 (1-year teaching experience); examiners 2 (6-year teaching experience) and 3 (1-year teaching experience) but there was no statistical difference in the mean marks given by examiners 1 and 2 for all the modules in TCE. In the OSPE, there was no significant difference in the mean marks given by all three examiners for all the modules, all being statistically significant at ( $P < 0.001$ ), suggesting that OSPE can be effectively used as an assessment tool with less experienced examiners by incorporating prevalidated checklists. Studies performed by Regehr *et al.*<sup>[26]</sup>

### Example of observed stations:

Palpate Liver/Spleen/Kidney (3 mks) (tick ✓ / X)										
Roll no:	1	2	3	4	5	6	7	8	9	10
Standing on right side of the patient (1/4)										
Introduction to the patient (1/4)										
Ensures the hands are warm (1/4)										
Position given to the patient (1/2)										
Ask the patient to take deep breaths (1/2)										
Places the hands correctly to examine liver/spleen/kidney (1)										
Thanking the patient (1/4)										
Total										

**Table 2:** Comparative study among examiners with different teaching experiences

		<i>Post hoc</i> Bonferroni test for multiple comparisons			
Teaching experience		TCE: A (1 + 2)	OSPE: A (1 + 2)	TCE: B (3 + 4)	OSPE: B (3 + 4)
Examiner 1 (35 yrs )	Examiner 2	0.523	1.000	0.071	1.000
	Examiner 3	<b>0.022*</b>	1.000	<b>0.001*</b>	1.000
Examiner 2 (6 yrs )	Examiner 1	0.523	1.000	0.071	1.000
	Examiner 3	<b>0.000*</b>	1.000	<b>0.000*</b>	1.000
Examiner 3 (1 yr)	Examiner 1	<b>0.022*</b>	1.000	<b>0.001*</b>	1.000
	Examiner 2	<b>0.000*</b>	1.000	<b>0.000*</b>	1.000

\* $P < 0.001$ .

have shown that global rating done by the experienced examiners is better and has a higher inter-station reliability, better construct validity, and concurrent validity than the checklists. Global rating done in TCE by subject experts is a more appropriate summative tool for assessing candidates on performance-based assessment.<sup>[17,18]</sup> However, use of global ratings mandates that only subject expert with adequate years of experience can be an examiner. In addition, Newbie<sup>[11]</sup> suggested that standardized checklists can be used for practical and technical skills stations whereas global rating scales can be used for stations concerned with diagnosis and communication skills concentrating on the affective domain of clinical competencies.<sup>[19]</sup> Another approach suggested was to use standardized checklists during early part of clinical training as a formative assessment and global ratings during final summative years.<sup>[17,20]</sup> We feel that a combination of OSPE with TCE would be a better parameter to do an overall assessment of the students. This needs a regular faculty development programs emphasizing on newer assessment tools like OSCE/OSPE to reduce the variability and improve objectivity. In addition, the use of a prevalidated checklist helped faculty review the common areas of weaknesses among students, which were discussed with them for improvement in their clinical skills. In addition, it can also highlight the areas of concern for the faculty to concentrate and improve their teaching-learning strategies. This use of OSCE/OSPE is well documented in literature,<sup>[7,22,25]</sup> which was well appreciated by many students. The present study faced some problems such as examiner's boredom and fatigue due to less interaction with students. We had selected criterion referencing, that is, 50% passing cutoff limit. The problem with criterion referencing is that poor performance in one station can be compensated by better performance in other.<sup>[17,23,24]</sup> It would be appropriate to decide the minimum pass marks and some important critical "must pass" stations to improve the validity of OSCE/OSPE. Also as only three examiners were the part of the study, the numbers of observation stations were restricted, reducing the accountability and validity of OSPE, straining the need for the continuous faculty development programs, and making them aware about the newer assessment tools in the medical field. The study also emphasized the need of a timely feedback from faculty to the students, which is an integral part of

OSPE making it more objective, valid, reliable and allowing variations in difficulty levels at various stations without compromising the uniformity.

## Conclusion

The study showed that OSPE can be effectively used as a assessment method with less experienced examiners by incorporating a prevalidated checklist. We emphasize the need of collaboration of OSPE with the traditional examination in both formative and summative assessments for a more global and comprehensive evaluation of students' skills and knowledge with the suggestion of development of a regular faculty program with a uniform international standardization of OSPE.

## Acknowledgments

The researchers are grateful to the Dean Dr. Geeta Niyogi; Dr. Lalit Mehta (director); Dr Avinash Supe, FAIMER; Dr. Rohini Bhadre (Department of Biochemistry); Dr S.S. Kowli; Mr Sinha (Department of PSM); Mrs. Judy; Mr. Sunil; Late. Mr Sharad; Mr Santosh (Department of Physiology); and the First MBBS Students 2010–2011 batch of K.J. Somaiya Medical College and Research Centre for their cooperation.

## References

1. Bartfay WJ, Rombough R, Howse E, LeBlanc R. The OSCE approach in nursing education: objective structured clinical examinations can be effective vehicles for nursing education and practice by promoting the mastery of clinical skills and decision-making in controlled and safe learning environments. *Canadian Nurse*. 2004;100(3):18–25.
2. Mandroli PS. OSCE based assessment of three commonly performed pediatric dentistry procedures. *Int J Clin Dental Sci*. 2010;1(1):64–69.
3. Dissanayake AS, Ali BA, Nayar U. The influence of the introduction of the objective structured practical examinations in physiology on student performance at King Faisal University Medical School. *Med Teach*. 1990;12(3–4):297–304.
4. Duffield KE, Spencer JA. A survey of medical students' views about the purpose and fairness of assessment. *Med Edu*. 2002;36:879–86.

5. Newble DI, Jaeger K. The effects of assessment and examinations on the learning of medical students. *Med Educ.* 1983;17:25–31.
6. Mohammed Al-Naami Y. Reliability, validity and feasibility of the objective structured clinical examination in assessing clinical skills of final year clerkship. *Saudi Med J.* 2008;(12):1802–07.
7. Ananthkrishnan N. Objective structured clinical/practical examination (OSCE/OSPE). *JPGM* 1993;39(2):82–4.
8. Miller G. The assessment o clinical skills/competence/performance. *Acad Med.* 65;1990(Suppl. 9):S63–S67.
9. Bijlani RL. Assessment of laboratory exercises in Physiology. *Med Educ.* 1981;15(4):216–21.
10. Edelstein DR, Ruder HJ. Assessment of clinical skills using video tapes of the complete medical interview and physical examination. *Med Teach.* 1990;12:155–62.
11. Newbie DI. The observed long case in clinical assessment. *Med Educ.* 1991;25:369–73.
12. Verma M, Singh T. Experiences with objective structure clinical examination (OSCE) as a tool for formative evaluation in pediatrics. *Indian Pediatr.* 1993;30:699–702.
13. Harden RM, Gleeson F. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ.* 1979;13:41–54.
14. Harden RM. What is an OSCE? *Med Teacher.* 1988;10:19–22.
15. Harden RM, Stevenson M, Wilson DW, Wilson GM. Assessment of clinical competencies using objective structured clinical examination. *Br Med J.* 1975;5955(1):447–51.
16. Friedman Ben-David M. The role of assessment in expanding professional horizons. *Med Teacher.* 2000;22(5):472–7.
17. Gupta P, Dewan P, Singh T. Objective structured clinical examination (OSCE) revisited. *Med Educ.* 2010(47):911–20.
18. Hasan S, Malik S, Hamad A, Khan H, BilalM. Conventional traditional practical examination (CPE/TDPE) verses objective structured practical examination (OSPE)/semi objective structured practical examination (SOSPE). *Pak J Physiol.* 2009;5(1): 58–64.
19. Nayar U, Malik SL, Bijlani RL. Objective structured practical examination: a new concept in assessment of laboratory exercises in preclinical Sciences. *Med Educ.* 1986;20:204–09.
20. Rahman N, Ferdousi S, Hoq N, Amin R, Kabir J. Evaluation of objective structured practical examination and traditional practical examination. *Med J.* 2007;16(1):7–11.
21. Roy V, Tekur U, Prabhu S. A comparative study of two evaluation techniques in pharmacology practicals: conventional practical examination versus objective structured practical examination. *Indian J Pharmacol.* 2004;36(6):385–9.
22. Downing SM. Reliability: on the reproducibility of assessment data. *Med Edu.* 2004;38:1006–12.
23. Hilliard RI, Tallett SE. The use of an objective structured clinical examination with postgraduate residents in pediatrics. *Arch Pediatr Adolesc Med.* 1998;152:74–8.
24. Joorabchi B. Objective structured clinical examination in a pediatric residency training program. *AJDC* 1991;145:757–62.
25. Carraccio C, Englander R. The objective structured clinical examination: a step in the direction of competency-based evaluation. *Arch Pediatr Adolesc Med.* 2000;154:736–41.
26. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med* 1998;73:993–7.

**How to cite this article:** Wani PD. Traditional clinical examination vs objective structured practical examination in human physiology: Examiner's bias. *Int J Med Sci Public Health* 2015;4:607-611

**Source of Support:** Nil, **Conflict of Interest:** None declared.