Experience of Objective Structured Practical Examination in Traditional Settings Of Human Physiology curriculum

Dr Pinaki Wani*,Dr Shobha Kini**, Dr Vrinda Dalvi ***
*Assistant Professor, **Assoc. Professor,*** Professor HOD, Department of Physiology , K.J.Somaiya Medical College , Sion 400022

Abstract: <u>Introduction:</u> An integral part of a medical curriculum is an appropriate assessment of the students' clinical competencies since assessment drives learning. A need of a more competence based assessment method led to introduction of Objective Structured Practical Examination (OSPE) which assesses the 'shows how' level of the Miller's pyramid of clinical competence as Traditional Clinical Examination (TCE) focuses on the "knows" and "knows how" aspects .The present study focuses on the experience of OSPE in term of the reliability and validity in comparison with traditional assessment method. <u>Methodology:</u> After the institutional ethical committee approval, a pilot study for comparing TCE with OSPE was conducted with a batch of 50 first MBBS students at K.J.Somaiya medical college in the department of Physiology for 4 consecutive days. Three examiners with teaching experience of 35, 6 and 1 year respectively conducted TCE followed by OSPE which had 10 stations assessing cognitive, psychomotor and affective domain for the same batch of 25 students in each. <u>Results:</u> OSPE had a good face validity and content validly as compared to TCE. Predictive validity using Pearson's correlation with the final year –end examination for TCE was 0.45 while for OSPE was 0.78 and reliability measured by internal consistency using Cronbach's alpha for TCE was 0.66 and for OSPE was 0.73 . The inter-station reliability measured affective and psychomotor domain in OSPE was 0.279 and 0.4 respectively while that for cognitive domain was 0.963. <u>Conclusion</u>: OSPE is a reliable and valid assessment tool provided it is more comprehensive and standardized. However it needs to be incorporated with traditional assessment for an overall evaluation of student's performance. [Wani P et al NJIRM 2013; 4(2) : 91-96]
Key Words: OSPE, TCE, reliability, validity

Author for correspondence: Dr. Pinaki Wani ,5, Deepa CHS, Sec – 2 A, Airoli , Navi Mumbai ,Maharashtra 400708, India. E-mail- wanipinaki@gmail.com

Introduction: The prime objective of a medical curriculum is to produce clinically competent medical practitioners. Hence an integral part of a medical curriculum is an appropriate assessment of the students' clinical competencies since assessment drives learning [1,2]. A framework for the development of clinical competence was described by Miller which outlines four levels of assessment: *"knows"," knows how", "shows how"* and *"does"*[.3.]The traditional practical examination in Physiology involves performing a clinical procedure followed by assessment based on global performance rather than the candidate's individual competency. In addition, the patient and examiner variability significantly affects the score. Traditional Clinical Examination (TCE) mainly focuses on the *"knows" and "knows how"* aspects, i.e. the base of the 'Miller's pyramid of competence' [4,5]. The objective structured clinical examination (OSCE) later extended to the practical examination (OSPE) described in 1975 and in greater detail in 1979 by Harden and his group from Dundee is defined as "approach to the assessment of clinical/ practical competence in which the components of competence are

assessed in a well planned or structured way with attention being paid to objectivity" [6,8]. The OSCE/OSPE conforms to the third *"shows how"* level of Miller's pyramid which focuses on assessment of performance of specific skills in a controlled setting [4,7,8]. This makes it particularly more relevant for the early stages of undergraduate curricula, where assessment comprises of compartmentalized exercises [9]. The OSCE/ OSPE has become a well-established method of assessment in medical education and is increasingly being used as a method of assessment in medical curricula [10]. An OSPE requires each student to demonstrate specific skills and behaviours in a simulated work environment with standardized patients. It typically consists of a circuit or series of short assessment tasks (stations), each of which is assessed by an examiner using a predetermined, objective marking scheme[10.]

In an attempt to make the assessment system more competence based and modern, OSPE was introduced as a pilot project for the first MBBS students during their second term in Human

Physiology. The present study focuses on the experience of the OPSE in the traditional settings of the assessment system in Physiology in India with its concern on the reliability and validity of both the assessment tools under study.

A good assessment tool must fulfil criteria's of objectivity, validity, reliability and feasibility. A reliable examination must be valid to be acceptable[13]. Validity of the tool is the extent to which it measures what it is supposed to measure and is the most important criteria for the assessment tool as stated by Newble [11,12] .Face validity refers to the extent to which an assessment tool measures what is considered important. Content validity refers to the extent to which an assessment tool covers the area of competence in questions [12,13,14]. Predictive validity refers to the association of the assessment tools with the outcome that can be predicted in future. Significant association would be an indication of a good predictive validity [13]. Concurrent validity refers to significant statistical association between the test results with another test or measure designed to assess the same attributes or behaviours done by comparing OSPE scores outcomes with the results of subjective ratings of the overall performance of practical skills for every participant immediately at the end of rating skills at each station. Agreement between OSPE test scores (objective) and the subjective rating of the overall performance would be an indication of a good OSPE's concurrent validity[13]. Reliability of the assessment tool refers to its consistency in discriminating examinees' performance and the reproducibility of assessment scores over time [13]. An acceptable reliability measured by Cronbach's alpha for high stakes (board certification), moderate stake (summative assessment), and low stakes (formative assessment) OSCE/OSPE are more than 0.9, 0.8 - 0.89, and 0.7 - 0.79, respectively [15]. Many factors influence reliability of the assessment tool including use of checklists as opposed to rating scales, standardized subjects to maximize reproducibility of individual station performances, minimum of 3- 4 hours of testing time, stations assessing hands-on clinical skills as opposed to stations using written items [16].

Objectives of the study: To report our experience of introducing a newer assessment tool in the form of OSPE and measuring the performance of the students in traditional and OSPE assessment tool along with comparing its reliability and validity with the existing traditional assessment tool available in Physiology.

Material and Methods: After the approval from Institutional Ethics committee for Research on Human Subjects, this pilot study was conducted in the Department of Physiology at the K.J. Somaiya Medical College and Research Centre, Mumbai, India.

A total of 50 first MBBS medical students (14 males and 36 females with the mean age of 18.6± 2.4 yrs) in their 2nd term in Human Physiology participated in the study after signing a written informed consent. All the students had ≥ 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 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 regarding the various assessment systems. The workshop had a two hours interactive session on OSCE/OSPE with demonstration on the formation of OSCE/OSPE stations and a practice session.

Following this, blueprint(desired clinical skills to be examined) of the structured checklist for observed and unobserved stations for 2 separate modules in abdominal system examination were prepared as per the Bloom's taxonomy along with examiner's and student's instruction manual and validated by the senior faculty members experienced in the field of medical education to ensure good content validity. The entire batch of 50 first MBBS students was divided in 2 batches of 25 each.

Each batch of 25 students was further divided in 3 batches with 8,8,9 students in each batch

examined by 3 examiners with teaching experiences of 35 years, 6 years and 1 year respectively who conducted TCE followed by OSPE for 4 consecutive practical days in Physiology. The examiners selected for the study had a basic training in the medical education innovations. 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 2 in abdominal system respectively and then there was a cross over done on the consecutive days. This was done to avoid any overlap and bias in learning by the students. Passing cut-off at 50 % based on criterion referencing was selected. In the traditional assessment method; each student performed a clinical skill which was followed by a bedside viva-voca on the same. The assessment of the student was global and not competence based. All the examiners were briefed about the general outline of the TCE format and marking scheme but no structured written format was followed. The questions were randomly asked based on the examiner's thoughts, perceptions, experience and mood. The entire TCE session ended in ninety to one- twenty minutes.

In the OSPE, for both the modules on abdominal system, there were 10 stations of 3-5 minutes each including 1 station on Communication Skills-affective , (1 mark), 2-3 procedure station-psychomotor domain, based on inspection, palpation and percussion of abdominal system (2-3 marks), 6 unobserved stations-cognitive domain, based on the procedural stations and MCQs (1-2 marks) + 1 rest station was arranged in Physiology practical lab in a clockwise manner. Students were oriented by an OSPE map and they moved from one station to another following the audible ring by the time keeper. An examiner appointed at procedural stations was provided with a pre-validated checklist to mark immediately according to the observed procedure. Answers given for the unobserved stations were checked as per the pre-validated model answers provided to all the examiners. The entire session ended in thirty to forty- five minutes.

The results were then submitted to the statistical committee for analysis using Microsoft Office 2007 Excel Software & SPSS software (Version 16.0). Face validity and content validity were judged by faculty and students participating in the study. Concurrent validity and predictive validity was measured. Internal consistency between TCE and OSPE along with inter-station reliability was measured by Cronbach's alpha.

Result: All the 50 first MBBS students participated in the study. They appeared for the TCE followed by OSPE for the same module. Their mean marks and range were calculated as depicted in graph 1, 2 respectively. Face and content validity of OSPE was analyzed. The marks obtained by the 50 students in both the assessment tools were compared in terms of their predictive validity using Pearson's correlation with the final year –end examination and reliability measured by internal consistency using Cronbach's alpha as described in Table 1.

Table:1 Paired Samples statistics

	TCE	OSPE
No. Of students	50	50
Mean score (max.marks = 30)	16.6400	15.0200
95% Confidence Interval for Mean- upper bound	15.2861	13.9559
Lower bound	17.9939	16.0841
Std. Deviation	4.76385	3.74417
Std Error	0.67371	0.52950
Minimum	7.50	6.70
Maximum	26.00	21.00
Range	18.50	14.30
Predictive validity using Pearson's correlation with the final examination	0.45	0.78
Internal consistency by Cronbach's alpha	0.66	0.73
Affective domain	-	0.279
Psychomotor domain	-	0.4
Cognitive domain	-	0.963

Graph1 : Traditional Clinical examination

Histogram

Mean =16.64
Std. Dev. =4.764
N =50
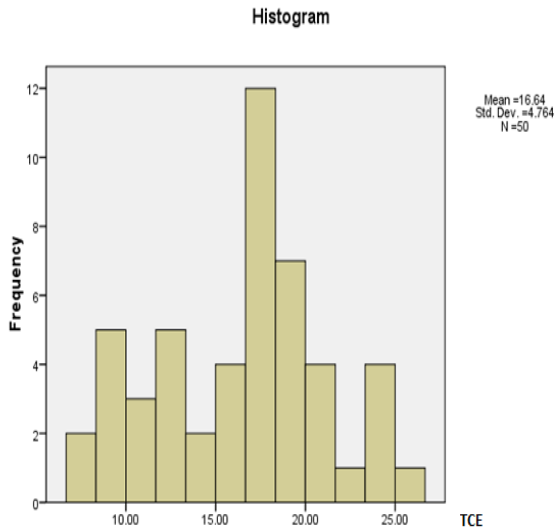
Graph 2: Objective Structured Practical Examination

Histogram

Mean =15.02
Std. Dev. =3.744
N =50
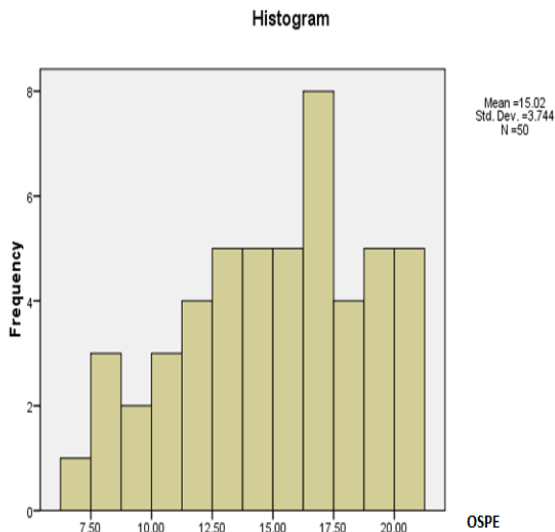
Discussion: With the Graduate Medical Education curriculum in the South-East Asia region moving its focus from knowledge to competency-based education, the assessment of students' clinical competencies becomes an integral part of the medical curriculum [17.] In an attempt to modernize the assessment system and make it more competence based, OSPE was introduced as a pilot project and compared with the traditional assessment tool.

Our study showed no statistical difference in the mean scores obtained by the students for both the assessment tools suggesting a good concurrent validity between the TCE and the OSPE. Also the range of marks in OSPE was limited as compared to TCE suggesting a more objective, focused approach of OSPE in assessing students. In addition, the structured format reduced the biasness and variability of markings by examiners in OSPE. However, similar study done by Rahman N et al and Roy V. showed significantly higher scores in OSPE compared to TCE suggesting different inclusion criteria for the examination [18,19]. OSPE had a better face validity compared to TCE as opined by both the students and the faculty members participating in the study since they felt that the OSPE measures more number of skills in depth which are relevant to the clinical practice. Similar finding were put forward by Newstead [20].

Content validity was ensured by extensive reviewing and blue printing of the curriculum done by the experts in the field which was specifically done with OSPE [12,13]. Predictive validity showed poor Pearson's correlation of TCE with the final examination (r= 0.45) while OSPE had a good correlation with the final year exams (r= 0.78) indicating that OSPE can be a good reference indicator as compared to TCE in judging the performance of the students over a period of time. However studies done by Brown, Cox ,Norman suggested that OSPE cannot assess competencies like holistic approach towards the patients, long term care towards patients pointing towards a poor predictive indicator [12].

Our study showed that the TCE had an internal consistency of 0.66 while the OPSE having 10 stations of 3 minutes duration had an internal consistency of 0.73 which is consistent with the acceptable reliability for the formative assessment ranging between 0.7-0.79 and better than TCE[15]. Studies done by Joorabchi [21] have shown that to achieve good reliability with internal consistency of >0.8 , OSPE must have more number of stations

assessing multiple skills with duration of OSPE between 4- 8 hours. The reliability of a 60 minutes OSPE session is as low as 0.54 and 0.69, respectively[3].

In our study, the total duration for the TCE ranged between 90-120 minutes while that for the OSPE ranged between 35-45 minutes for each session held on 4 consecutive days. But it is impractical to conduct an OSPE of more than 60 minutes duration with many stations due to examiners and students fatigability.

This would affect the student's performance and examiner's reluctance due to work overload. The inter-station reliability measured by Cronbach's alpha for affective and psychomotor domain in OSPE was 0.279 and 0.4 respectively while that for cognitive domain was 0.963 straining the need for a more comprehensive training and standardization for judging clinical skills and attitude. In addition, all the students were exposed to similar competency based stations for the same duration with the pre-validated checklist for marking which improved the reliability of OSPE in comparison with TCE where majority of the questions asked by the examiners were based on their own perceptions, thinking and mood as confirmed by the study by Hilliard and Tallett [22] who performed OSCE with 5 stations and achieved a Cronbach's alpha score of 0.69 by virtue of homogeneity of tasks at different stations and standardized patients .

Our study enrolled trained standardized subjects who could enact and reply to the examinee's questions appropriately reducing the patient variability and improving the reliability of OSPE. But as mentioned by Verma and Singh [23], OSPE assesses only specific clinical skills and hence it needs to be collaborated with other assessment methods like TCE with global approach for the comprehensive assessment of the students. The study had limitations as the numbers of stations were restricted to 10 with only 3 observed stations and 6 unobserved stations due to less faculty participation and lack of time to incorporate medical educational projects in the hectic schedule of first and second term affecting the reliability and the validity of the tool.

In addition, the low reliability in the affective and psychomotor domains emphasises the need of a more comprehensive teaching learning effort. Also it highlights the need of forming multiple OSPE stations with valid, structured and standardized checklist to assess wide range of affective, psychomotor and cognitive skills along with a continuous faculty development programmes to acquaint them with innovations in medical education and technology.

Conclusion: With the focus of medical education in India shifting from knowledge based to a more competency based tool, our study showed that OSPE is a more reliable and valid assessment tool in comparison to TCE which is required for the formative assessment but it cannot solely replace the existing assessment system and it needs to be incorporated with TCE for a overall judgement of the students performance.

References:
1. Norman GR. Current issues in the assessment of clinical competencies. In: Hart IR, Harden RM editors: Further developments in assessing clinical competencies. Montreal Que:Can- Heal Publication Inc ; 1987.p. 136-147
2. Newble Dl, Entwistle NJ. Learning styles and approaches: implications for medical education .Med Education 1986;20:162-75
3. Miller, G.. The assessment o clinical skills / competence / performance. *Academic Medicine 65, 1990 (Suppl.9), S63-S67*
4. Piyush Gupta, Pooja Dewan , Tejinder Singh. Objective Structured Clinical Examination (OSCE) Revisited. Medical education 2010(47):911-920

5. Smith LJ, Price DA, Houston IB. Objective structured clinical examination compared with other forms of student assessment. *Arch Dis Child* 1984;59:1173-6.

6. Harden, R. M.. What is an OSCE? Medical Teacher 1988 :10(1), 9-22.

7. Hamdy, H. Blueprinting for the assessment of health care professionals. The Clinical Teacher 2006 :3, 175-179.

8. Harden, R. M., & Gleeson, F. A. Assessment of clinical competence using an objective structured clinical examination (OSCE). Medical Education 1979:13, 41-54.

9. Friedman Ben-David, M. The role of assessment in expanding professional horizons. Medical Teacher 2000:22(5), 472-477.

10. Bartfay, W. J., 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. The Canadian Nurse 2004:100(3), 18-25.

11. Newble Dl. Assessing clinical competence at the undergraduate level. Med Edu 1992:26:504-11

12. A Barman . Critiques on the Objective Structured Clinical examination . Ann Acad Med Singapore 2005;34:478-82

13. Mohammed Y. Al – Naami. Reliability , validity and feasibility of the Objective Structured Clinical examination in assessing clinical skills of final year clerkship. Saudi Med J 2008;(12):1802-1807

14. Abdullah J. Rekany, Samim A. Al-Dabbagh. Validity And Reliability Of Osce In Evaluating Practicalperformance Skills Of Interns In Emergency Medicine Duhok Med J 2010;4(2):21-29

15. Downing SM. Reliability: on the reproducibility of assessment data. Med Edu 2004;38:1006-12

16. Van der Vleuten CPM, Swanson DB. Assessment of clinical skills with standardized patients: state of the art. *Teach Learn Med.* 1990;2:58-76..

17. http://psgimsr.in/faculty%20

18. 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.

19. Hasan S, Malik S, Hamad A, Khan H, Bilal M. Conventional / traditional practical examination versus Objective structured practical examination (OSPE)/Semi Objective structured practical examination (SOSPE) Pak J Physiol 2009;5(1)

20. Newstead, S. E. The use of examinations in the assessment of psychology students. Psychology Teaching Review1992; 1(1), 22-33

21. Joorabchi B. Objective structured clinical examination in a pediatric residency training program. *AJDC.* 1991;145:757-762.

22. Hilliard RI, Tallett SE. The use of an objective structured clinical examination with postgraduate residents in pediatrics. *Arch Pediatr Adolesc Med.* 1998;152:74-78.

23. Verma M, Singh T. Experiences with objective structured clinical examination (OSCE)as a tool for formative evaluation in Pediatrics . Indian Pediatrics 1993;30:699-702