Comparing Students' Performance in Final Year OSCE: Station Type and Clinical Skills Assessed

Abstract

Background: In our instituiton, objective structured clinical examination (OSCE) is a component of the Final MBBS Examination. The purpose of this paper is to analyse students' performance in OSCE. The two objectives were to compare students' performance: (i) in interactive and non-interactive stations, and (ii) in the six clinical skills assessed.

Method: Data for this study were obtained from the Final MBBS Examination 2012 (n=185). For the 16-stations OSCE, nine were interactive and seven were non-interactive. For interactive stations, both checklists and global ratings were used for scoring. For non-interactive stations, only checklists were used. Each station's score sheet comprised a detailed checklist of items examined (total=10marks). Global rating was also included for the examiner to indicate the global assessment for the station. Retrospective analysis of data was conducted using SPSS. Means for interactive and non-interactive stations were computed and compared. Means for the six skills assessed were also computed and compared.

Results: Means for interactive and non-interactive stations were respectively 6.16±0.97 and 5.77±1.09. Paired sample t-test showed students' performed significantly better in interactive stations, at $p<0.001$. Means for history taking, physical examination, communication skill, clinical reasoning skill (CRS), procedural skill and professionalism were respectively 6.25±1.29, 6.39±1.36, 6.34±0.98, 5.86±0.99, 6.59±1.08 and 6.28±1.02. Repeated measures ANOVA showed significant differences in students' performance in the six clinical skills assessed, at $p<0.001$.

Conclusion: Students performed significantly better in interactive compared to non-interactive stations. Procedural skills appeared to be the strongest while CRS was the weakest among the six clinical skills assessed.

Comparing Students' Performance in Final Year OSCE: Station Type and Clinical Skills Assessed

University of Malaya

Sim JH, Yang F, Anushya V , Azura M, Vadivelu J
Faculty of Medicine
University of Malaya
Kuala Lumpur, Malaysia

Introduction

In our institution, objective structured clinical examination (OSCE) is a component of the Final MBBS Examination

2

Introduction

- The purpose of this study was to analyse students' performance in Final Year OSCE
- The two objectives were to compare students' performance in:

  (i) interactive and non-interactive stations, &

  (ii) the 6 different clinical skills assessed

Introduction

The concept of "clinical skill" is not clearly defined in the literature (Michels, Evans & Blok 2012)

Operational definitions:

- "Interactive station"- a station where there is some form of interaction (between candidate and examiner and / standaridised patient / mannequin)
- "Non-interactive station"- a station where there is no direct observation and assessment
- The 6 clinical skills assessed were HT, PE, Comm, CR, Proc, Prof)

Method

- Data for this study were obtained from the Final MBBS Examination 2012 (n=185)
- 16 work stations & 1 rest station
- 5 minutes per station
- 3 parallel tracks/circuits
- 4 rounds

5

Method

- 9 interactive & 7 non-interactive stations
- interactive stations: both checklists & global ratings for scoring
- non-interactive stations: checklists only
- station's score sheet comprised a detailed checklist of items examined (total=10 marks)
- Global rating - for the examiner to indicate the global assessment for the station

6

Method

Measures taken to increase validity & reliability

- Content validity was established by blueprinting
- For quality assurance, question vetting was conducted (at department & faculty level)
- 16 stations from 11 clinical departments (ensure wider sampling across subject areas & skills)
- Stations were reviewed & field-tested
- Training of examiners + structured marking schedules
- Training of standardised patients

7

Method

- Raw score for each station (n=185) was obtained
- Retrospective analysis of data using SPSS
- Cronbach alpha for the 16 stations was computed
- Means for interactive & non-interactive stations were computed and compared using paired-sample t-tests
- Means for the 6 skills assessed were computed and compared using repeated measures ANOVA (within-subjects design)
- An alpha level of 0.05 was set for all the statistical tests

8

Results

- Reliability analysis reported an alpha value of 0.68 (n=185)
- Acceptable internal consistency or reliability for the 16-station OSCE

9

Results

Table 1: Paired-sample t-test (Descriptive statistics)

Station Type	N	Mean	S.D.
Interactive	185	6.16	0.97
Non-Interactive	185	5.77	1.09

- Mean for interactive stations was higher compared to non-interactive stations

10

Results

Table 2: Paired sample t-test of interactive and non-interactive mean scores (n=185)

Pair	Paired Differences			t	df	Sig.
	Mean	S.D.	Std. Error Mean			
Interactive-Non-Interactive	0.39	0.962	0.071	5.573	184	0.000

- Paired sample t-test showed students performed significantly better in interactive stations, [$t(184)=5.573$, $p<0.001$]

11

Results

Table 3: Repeated measures ANOVA (Descriptive statistics)

Clinical Skills Assessed	N	Mean	S.D.
History Taking	185	6.25	1.29
Physical Examination	185	6.39	1.36
Communication Skill	185	6.34	0.98
Clinical Reasoning Skill (CRS)	185	5.86	0.99
Procedural Skill	185	6.59	1.08
Professionalism	185	6.28	1.02

- Mean for CRS was the lowest while mean for procedural skills was the highest among the 6 clinical skills assessed

12

Results

Table 4: Mauchly's Test of Sphericity[a]

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Skills assessed	0.050	546.824	14	**0.000**	**0.596**	0.607	0.200

a. Design: Intercept
Within Subjects Design: Skills assessed

b. May be used to adjust the degrees of freedom for the average tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table (Table 5)

- Mauchly's Test of Sphericity is significant
- Adjustment of df for the test in tests of within-subjects effects need to be done

13

Results

Table 5: Overall analysis of variance-Tests of Within-Subjects Effects

Source of variation	Sum of squares	df	Mean square	F	Sig.
Skills					
Sphericity Assumed	51.876	5	10.375	20.253	0.000
Greenhouse-Geisser	51.876	**2.980**	17.408	**20.253**	0.000
Huynh-Feldt	51.876	3.035	17.095	20.253	0.000
Lower-bound	51.876	1.000	51.876	20.253	0.000
Error (Skills)					
Sphericity Assumed	471.283	920	0.512		
Greenhouse-Geisser	471.283	**548.332**	0.859		
Huynh-Feldt	471.283	558.355	0.844		
Lower-bound	471.283	184.000	2.561		

- Based on the new df, there was also a significant difference among the 6 clinical skills assessed, [$F(2.980, 548.332)=20.253$, $p<0.001$]

14

Results (Pairwise multiple comparisons)

Pairs of skills compared	Mean difference	Std. Error	Sig.[a]
History-Examination	-0.143	0.112	1.000
History-Communication	-0.087	0.074	1.000
History-Clinical Reasoning Skills	**0.380***	0.091	**0.001**
History-Procedural Skills	**-0.340***	0.096	**0.008**
History-Professionalism	-0.031	0.072	1.000
Examination-Communication	0.055	0.077	1.000
Examination-Clinical Reasoning Skills	**0.523***	**0.089**	**0.000**
Examination-Procedural Skills	-0.197	0.089	0.425
Examination-Professionalism	0.112	0.075	1.000
Communication-Clinical Reasoning Skills	**0.467***	**0.049**	**0.000**
Communication-Procedural Skills	**-0.252***	0.039	**0.000**
Communication-Professionalism	0.056	0.027	0.531
Clinical Reasoning Skills-Procedural Skills	**-0.719***	**0.062**	**0.000**
Clinical Reasoning Skills-Professionalism	**-0.411***	**0.061**	**0.000**
Procedural Skills-Professionalism	**0.309***	0.054	**0.000**

15

Results

Table 6: Multiple comparisons (pairwise)

Pairs of skills compared	Mean difference	Std. Error	Sig.[a]
CRS-History	-0.380*	0.091	0.001
CRS-Examination	-0.523*	0.089	0.000
CRS-Communication	-0.467*	0.049	0.000
CRS-Procedural Skills	-0.719*	0.062	0.000
CRS-Professionalism	-0.411*	0.061	0.000
Procedural Skills-History	0.340*	0.096	0.008
Procedural Skills-Communication	0.252*	0.039	0.000
Procedural Skills-Professionalism	0.309*	0.054	0.000

*Mean difference is significant at the 0.05 level

a. Adjustment for multiple comparisons: Bonferroni correction

16

Conclusion & Take Home Message

- Students performed significantly better in interactive (M=6.16) compared to non-interactive stations (M=5.77) [t(184)=5.573, p<0.001]
- There was a significant difference among the six clinical skills assessed [F(2.980, 548.332)=20.253, p<0.001]
- CRS (M=5.86) appeared to be the weakest skill while procedural skills (M=6.59) was the strongest, among the skills assessed
- Students' unsatisfactory performance in CRS needs to be addressed

17

THANK YOU

18