Dwivedi NR, Vijayashankar NP, Hansda M, Dubey AK, Nwachukwu F, Curran V, Jillwin J. Comparing Standard Setting Methods for Objective Structured Clinical Examinations in a Caribbean Medical School. J Med Educ Curric Dev. 2020;8:2382120520981992

The authors regret a typographical error in their statistics in the 3rd paragraph on page 6 (4th paragraph of the Relative Method section within the Results). The authors would like to update the following sentence:

“The effect size was calculated using Cohen’s “d” values for the pooled standard deviation of 5.57.”

To

“The effect size was calculated using Cohen’s “d” values as |70-M|/SD, where 70 is the given traditional standard, M and SD are the observed mean and the estimated standard deviation.

The paragraph now reads as follows:

“The effect size was calculated using Cohen’s “d” values as |70-M|/SD, where 70 is the given traditional standard, M and SD are the observed mean and the estimated standard deviation. A One-sample t-test shows that the CVS mean (M = 64.40, SD = 7.13) was significantly below the traditional standard of 70, t (64) = −6.33, P < .001. The effect size was “large” d = 0.84. Likewise, a One-sample t-test shows that the RS mean (M = 67.2, SD = 6.10) was significantly below the traditional standard of 70, t (54) = −3.32, P < .002. The effect size was “medium” d = 0.59. The GIS mean (M = 69.6, SD = 8.26) was not different than the traditional standard of 70, t (52) = −.31, P > .02. The effect size was “small” d = 0.05. For these multiple comparisons, alpha was adjusted from .05 to .02 to reduce the risk of a Type 1 error.”

The authors would also like to add further detail in the Methods section regarding the validity of the grading criteria and checklists.

Methods

“There is no data collected to evaluate inter-rater reliability for this study because each trained and qualified examiner took responsibility for single station of each organ system (there are no mix-up of different examiners per station), thus we did not evaluate the inter-rater reliability. Each examiner was trained and tested for their qualification as an examiner.

To ensure consistency and fairness of scores, all the faculty were trained gradually to conduct OSCE’s during which they were clearly informed about the objectives, outcomes, roles and responsibilities and were allowed to shadow and observe.

All the OSCE examiners were MD’s by qualification. The new examiners/faculty (trainee) were trained systematically via workshops before they became examiners. This involved:

1. The first step consisted of orientation to the new faculty (trainee) in which knowledge, attitudes and skills are conveyed by the experienced faculty/examiner (trainer). This contained information about: OSCE as an examination, the process on the examination day, the role of examiners and sources of bias. Rubrics and previous recorded performances of students are included to ensure proper understanding of the information.
2. New faculty/examiners were then required perform a mock grading of the recorded video performance of student with the help of the rubrics.
3. The third step was shadowing of the experienced faculty/examiner (trainer) during the live OSCE.
4. During live OSCE, the experienced faculty/examiner (trainer) observed the new faculty/examiner and provided the feedback to them. This step was repeated as required.
5. The steps 2 to 4 is repeated before every OSCE for a minimum of 3 organ system live OSCE’s or until the trainer decided.

For the validity of the checklists, in this paper, we only state that the station checklists were reviewed and validated by all members of faculty involved in the study.
To begin with, the checklists that were being used for assessment of OSCE’s at Xavier University was used for this study. These checklists were developed initially by the subject expert and were gradually standardized and updated based on the feedback over a period of 9-10 years.

The feedback was sought by the clinical chairs, subject experts, based on the student performances, and graduate feedbacks. This was also modified based on the evolving needs of the medical curriculum. During the designing process of the study, a modified Delphi method was used to design the stations and adopting the checklists.

In this study, the aim and objective of the study was determined first, followed by designing the station and choosing the rubric or checklists from the pre-existing rubrics that were being used for OSCE’s.

Once the stations and checklists were finalized, the content validity was determined by all the 5 examiners who reviewed the checklists and provided their opinions and this was discussed. After these discussions, the final check list that was approved by all the 5 examiners was chosen in the study.”