Standard setting of objective structured practical examination by modified Angoff method: A pilot study

M. GANESH KAMATH, VINOD PALLATH, K. RAMNARAYAN, ASHA KAMATH, SHARMILA TORKE, JAMES GONSALVES

ABSTRACT

Background. The undergraduate curriculum at our institution is divided system-wise into four blocks, each block ending with theory and objective structured practical examination (OSPE). The OSPE in Physiology consists of 12 stations, and a conventional minimum score to qualify is 50%. We aimed to incorporate standard setting using modified Angoff method in OSPE to differentiate the competent from the non-competent student and to explore the possibility of introducing standard setting in Physiology OSPE at our institution.

Methods. Experts rated the OSPE using the modified Angoff method to obtain the standard set cut-off. We assessed the OSPE marks of 110 first year medical students. Chi-square test was used to compare the number of students who scored less than standard set cut-off and conventional cut-off; correlation coefficient was used to assess the relation between OSPE and theory marks in both blocks. Feedback was obtained from the experts.

Results. The standard set was 62% and 67% for blocks II and III, respectively. The use of standard set cut-off resulted in 16.3% (n=18) and 22.7% (n=25) students being declared unsuccessful in blocks II and III, respectively. Comparison between the number, who scored less than standard set and conventional cut-off was statistically significant (p=0.001). The correlation coefficient was 0.65 (p=0.003) and 0.52 (p<0.001) in blocks II and III, respectively. The experts welcomed the idea of standard setting.

Conclusion. Standard setting helped in differentiating the competent from the non-competent student, indicating that standard setting enhances the quality of OSPE as an assessment tool and the examinee-centred methods (borderline-group, contrasting groups and Hofstee's methods). Test-centred methods are those where the test items are reviewed by experts to pass a judgement on the test items as 'just an adequate' level of performance. The experts identify an actual group (not a hypothetical) and make judgements on the examinees.4

In the Angoff method, the experts make judgements on how hypothetical borderline candidates will perform on each item/question and obtain a cut-off score. The modified Angoff method has additional steps wherein the experts are provided with data on actual performance to understand the difficulty of test items and relook at the scores provided. The Ehel method requires the experts to categorize the test items based on the difficulty level and then decide on the proportion of items in each of the categories, the hypothetical borderline students can answer appropriately. The number of items and the number answerable by the hypothetical group in each category is multiplied to get a standard score. The Nadecky method is specifically designed for multiple-choice questions, requires the experts to decide how many distractors to have in a test item, a minimally competent student will be able to identify as inappropriate/wrong. In the Jaeger method, the experts rate each item, based on whether an examinee can answer correctly or incorrectly. This method is followed by another group of experts. After which, each of the experts can relook their assessment based on the others’ assessments. This method emphasizes more on passing examinees, than hypothetical borderline students.4

In the borderline group method, the experts identify an actual borderline group instead of a hypothetical group. The scores given by the experts are used to obtain a ‘median score’, which is then used as the passing score. This contrasts with the group approach in that it requires the experts to divide the examinees into those who are non-competent and competent. This assessment is based