Frontiers of Nursing • 12(4) • 2025
DOI: 10.2478/FON-2025-0056

Frontiers of **Nursing**

Comparing the passing rates of nursing students participating in the objective structured clinical examination using fixed score, Angoff, yes/no Angoff, and 3-level Angoff methods†

Original article

Fatemeh Maghsoodi[a], Mohammadreza Yazdankhahfard[a,*,#], Shahnaz Pouladi[a], Kamran Mirzaei[b], Amin Beigzadeh[c,*,#]

[a]*Department of Nursing, Faculty of Nursing and Midwifery, Bushehr University of Medical Sciences, Bushehr, Iran*
[b]*Department of Social Medicine, Faculty of Medicine, Shiraz University of Medical Sciences, Shiraz, Fars, Iran*
[c]*Department of Medical Education, Education Development Center, Sirjan School of Medical Sciences, Sirjan, Iran*

Received: 25 November 2024; Accepted: 26 January 2025; Published: 20 December 2025

Abstract: **Objectives:** To compare the frequency of passing scores of nursing students participating in objective structured clinical examination (OSCE) by 4 methods of fixed score, Angoff, yes/no Angoff, and 3-level Angoff. The OSCE is a valid and reliable tool to evaluate the competency of nursing students. To have a fair judgment in OSCE, it is important to determine the standard setting.
Methods: In this descriptive-analytical study, which was conducted in 2022, the 7-station OSCE scores and the passing rates of students were compared based on fixed score, Angoff, yes/no Angoff, and the 3-level Angoff by a panel of experts. The research population included nursing teachers (n = 13) and nursing students (n = 65). Due to the small number of participants, the number of samples in each of these 2 groups was equal to the total number of people in the research population.
Results: Findings revealed that of the 13 nursing teachers who participated in the study, the majority were female and had a master's degree. Among the 65 students taking part in OSCE, the majority were males. Results showed that the passing rate varied depending on the type of OSCE station and the standard-setting methods used. Differences in some cases were statistically significant ($P < 0.001$).
Conclusions: Findings suggest that the frequency of passing can vary depending on the method used. It is advisable to consider the average score from multiple methods to accurately determine the passing rate and achieve a more accurate acceptance rate.

Keywords: *nursing • OSCE • standard setting • assessment • reliability*

†This project was supported by vice-chancellery for research of Bushehr University and the National Agency for Strategic Research in Medical Sciences Education (NASR) of Tehran (Code: 4010071).

#Mohammadreza Yazdankhahfard and Amin Beigzadeh are co-corresponding authors.

How to cite this article: Maghsoodi F, Yazdankhahfard M, Pouladi S, et al. Comparing the passing rates of nursing students participating in the Objective Structured Clinical Examination using fixed score, Angoff, yes/no Angoff, and 3-level Angoff methods. *Front Nurs*. 2025;4:511–518.

*Corresponding author.
E-mail: myazdankhahfard@yahoo.com (M. Yazdankhahfard); beigzadeh.amin@gmail.com (A. Beigzadeh).

1. Introduction

Assessment, a key component of competency-based medical education, is widely recognized as a valuable method to verify learning achievements. It is not only a tool for verification of learning outcomes but also a driver of motivation, directly influencing students' attitudes and behaviors toward their learning. The nature of assessment, particularly in medical and nursing education, demands that its results be justifiable to multiple stakeholders, including students, evaluators, institutions, and the wider professional community.[1] The concept of fairness and transparency in assessment is vital to maintaining the integrity of the educational process and ensuring that students' achievements are accurately reflected.[2]

In recent decades, there have been significant advances in our understanding and use of different methods of assessment, especially workplace-based assessments.[3] These assessments, which closely mirror real-world clinical tasks, have become integral in bridging the gap between theory and practice.[4] This shift emphasizes the need for assessments that are not only reliable and valid but also applicable to the specific demands of the professional practice that students are preparing to enter.

Nursing students need to acquire not only nursing knowledge but also clinical skills during their studies before graduation. Evidence shows that nursing students must be assessed with different modes of assessment, and the objective structured clinical examination (OSCE) is of utmost importance accordingly.[4] Among the available assessment methods to evaluate students' clinical competence, the OSCE is a valid and reliable tool to discriminate between students in terms of mastery over clinical competence. In the OSCE, the elements of competency are assessed in a structured way, with a focus on maintaining objectivity.[5] Nursing curriculum has also put a considerable emphasis on OSCE implementation for nursing students.[6]

At our University, apart from the formative assessment practices, the OSCE, as a high-stakes examination, continues to be dominant in assessing nursing students' clinical skills.[7] The OSCE allows for a systematic and standardized approach to assess clinical skills, ensuring that students are tested on specific competencies through a series of stations, each designed to evaluate different aspects of their clinical performance.[7] Evidence consistently supports the validity and reliability of OSCEs as a tool for differentiating levels of clinical competence among nursing students.[8] Moreover, nursing education, which emphasizes both theoretical knowledge and practical, hands-on clinical skills, has increasingly integrated OSCEs as a key component in the curriculum.[9] The use of OSCEs enables a multifaceted evaluation of nursing students, ranging from technical skills to communication and critical thinking skills.[7,8] However, the growing reliance on this method also raises important considerations regarding its implementation and its capacity to reflect the complexity of nursing practice truly.

In high-stakes examinations, it is highly important to design effective assessments as well as determine the standard setting to differentiate between candidates who have reached the required level of proficiency and those who have not attained the essential professional skills. Standard setting in OSCEs is pivotal for ensuring fairness and consistency in assessing clinical skills, and it defines the minimum competence level required to pass.[10] This process incorporates various methods, among the most commonly used methods are the mean borderline group method, borderline regression, and the Angoff method. Based on the mean borderline group method, examiner opinions provide an overall assessment of examinees' performance in each station, separate from the checklist score. The average score of all borderline examinees is then used as the passing threshold for that particular station. By calculating the average borderline score for each station, the overall pass mark for the entire OSCE can be determined.[11] This approach is seen as a relatively practical and straightforward method but has been critiqued for its subjective nature and potential for inconsistency.[12,13] Additionally, the borderline regression method involves using linear regression to predict total OSCE scores based on global ratings. Pass marks are determined by substituting the scores of borderline candidates into the regression equation for each OSCE.[14] However, critics argue that it may fail to capture the nuanced, multifaceted nature of clinical competence, as it assumes that performance across stations is linear, which may not always be the case.[15] In the modified Angoff method, judges examine each question following the identification of a marginal candidate and determine whether the candidate would answer correctly.[16] This method is considered one of the acceptable methods of setting different pass marks.[17] However, it has been criticized for its time-consuming nature, complex organization, expense, and the significant variation in judgments that can arise among raters.[18] All in all, research findings accentuate that setting consistent standards for evaluating students' performance systematically and objectively is essential in medical education.[15,19]

A recent study by Park[20] found that the Angoff method was more reliable than the yes/no Angoff method in estimating the cut score of the Korean Medical Licensing

Examination (KMLE). However, the yes/no Angoff method with a 60% probability for deciding "yes" can also be used as a substitute for the Angoff method in estimating the cut score of the KMLE. In a study aimed to compare a rational and an empirical standard-setting procedure conducted by Kramer et al.[18], the borderline regression method provided a more credible and reliable standard for an OSCE than a modified Angoff procedure. The researchers concluded that a reality check improves the credibility of the Angoff procedure but does not improve its reliability. A study by Hejri et al.[21] assessed the reliability of borderline regression method when the pass-fail standard in an OSCE was calculated by averaging the borderline regression method standards obtained for each station separately and found that the root mean square error of the standard was very small indicating that borderline regression method is a reliable method of setting standard for OSCE, which has the advantage of providing data for quality assurance. A study conducted by Dwivedi et al.[22] at Xavier University School of Medicine in 2020 compared 4 different standard-setting methods with OSCEs. This study indicated that the traditional standard-setting method resulted in the highest failure rates. In contrast, the modified Angoff method and borderline regression method demonstrated consistent reliability and practical suitability, with an average of these methods providing an acceptable cutoff score. In a similar vein, Dwyer et al.[17] at a Canadian University in 2016 used the Angoff method to set pass marks. This was then compared to those determined by the Borderline Group and Borderline Regression methods. This study affirmed the modified Angoff method as a suitable approach for setting differentiated pass marks for varying levels of expertise.[17]

Based on recent studies on standard setting, research findings indicate that various methods of setting standards can result in varying levels of acceptability.[23] To the best of our knowledge, there is a paucity of research in the field of Nursing focusing on various approaches to standard setting in nursing OSCEs. Therefore, the purpose of this study was to fill this gap by comparing the frequency of passing scores of nursing students participating in OSCE by 4 methods of fixed score, Angoff, yes/no Angoff, and 3-level Angoff.

2. Methods

2.1. Study design

This descriptive, analytical cross-sectional study was undertaken on 13 nursing teachers who were members of the OSCE committee and 65 nursing students in the 8th semester in 2022. Participants entered the study based on census sampling.

2.2. Inclusion and exclusion criteria

The criteria for teachers encompassed: (a) being affiliated with the nursing department as either faculty or non-faculty, (b) having experience in OSCE implementation, (c) having experience in teaching or internship with nursing students, and (d) participating in a standard setting workshop. A total of 13 participants, consisting of 10 females and 3 males, met these criteria. Among them, 10 participants were married, and 3 participants were single. In terms of educational background, 9 participants held a Master of Science degree, and 4 participants held a PhD certificate.

The entry criteria for students included: (a) enrollment in the 8th semester of nursing, (b) successful completion of all theoretical, practical, and internship courses, and (c) participation in the OSCE examination. The nursing OSCE exam consisted of 7 stations where students had 5 minutes at each station to complete the assigned task. Due to the limited number of nursing students in the 8th semester (N = 65), all students met the criteria to enter the study (excluded students n = 0). Participants were recruited using the census method.

2.3. Study approach

In this study, the nursing OSCE was conducted using 4 different methods of standard setting. These methods included: (1) fixed score, (2) the Angof method, (3) the yes/no Angof method, and (4) the 3-level Angof method. The following section describes each standard-setting method. For the fixed score, due to the clinical nature of the OSCE exam, the cutoff of 12 (60 out of 100) was considered the minimum score required to pass the OSCE. To determine the standard setting for Angoff, yes/no Angoff, and the 3-level Angoff methods, after developing the OSCE stations and preparing the relevant checklists by the OSCE committee, we invited nursing teachers to participate in the standard-setting workshop. The workshop lasted for 1 day, participants were instructed on standard setting, and their questions were addressed accordingly. During the workshop, a key topic of discussion was the classification of examiner performance levels as excellent, satisfactory, borderline, and failed. In test-based assessments, the primary objective is to determine the passing score for test questions at various stations or scenarios of the OSCE. Each nursing teacher was tasked with evaluating and ranking the 7 stations individually, considering the probability of passing by a borderline student based on different standard-setting methods. In this regard, nursing teachers were asked to score based on the following values. The standard setting for the Angoff method was a value between 0% and 100%. Yes/no Angoff was as 100 points for

"yes" and 0 points for "no." Finally, for the 3-level Angoff, answers were categorized as "yes," "no" and "maybe" (100 points for the answer "yes," 50 points for "maybe," and 0 points for the answer "no").

In this manner, each teacher stated a numerical rating for each station, with the mean of these ratings being considered as the standard for that particular station (pass or fail cut-off point for that station). Besides, the mean standard of different stations was considered as the standard of the whole OSCE test. Then, after the implementation of the OSCE, the scores obtained by each student were compared with the cut-off point of each station, and the students were classified into 2 groups, pass and fail. In the following, the frequency of accepted and rejected students was compared with each other based on 2 evaluation methods. It is worth mentioning that there were no confounding variables in this study. To analyze the data, descriptive statistics (frequency and percentage) were used. In addition, the chi-squared test was used for inferential statistics. The data were analyzed by using SPSS software (IBM, Chicago, IL, USA) version 24.

3. Results

Findings revealed that the majority of nursing teachers (n = 13) who took part in standard setting were women (76.9%), married (76.90%) and had a master's degree (69.20%). Furthermore, the majority of the 65 nursing students who took part in the OSCE were male, accounting for 50.80% of the participants. The age range of the students varied from a minimum of 22 years to a maximum of 49 years.

The data presented in Tables 1–6 show the distribution of student scores and the frequency of passing across different stations using various standard-setting methods.

As can be seen from Table 1, in pharmaceutical care, vaccination, gavage, and IV therapy stations, the passing percentage in OSCE using the Angoff score method was significantly higher than the passing percentage using the fixed score method. There were no significant differences observed in other stations.

Furthermore, as depicted in Table 2, the frequency of passing in OSCE using the fixed score method at vaccination, gavage, and IV therapy stations showed a significant difference compared to the frequency of passing obtained using the yes/no Angoff method. In this regard, the fixed score method resulted in a higher acceptance rate at the vaccination and gavage stations, while the yes/no Angoff method yielded a higher acceptance rate at the IV therapy station.

According to the data presented in Table 3, the frequency of passing scores in the OSCE at the gavage and IV therapy stations was significantly higher using the 3-level Angoff method compared to the fixed score method.

As shown in Table 4, in the gavage station, the frequency of achieving a passing score in OSCE through the Angoff method was significantly higher than the passing score obtained through the yes/no Angoff method. Conversely, at the IV therapy station, the frequency of passing scores using the yes/no Angoff method exceeded those achieved through the Angoff method. These distinctions were found to be statistically significant.

Table 5 indicates a notable difference in passing score frequencies between the Angoff method and the 3-level Angoff method for gavage and IV therapy stations. Specifically, the passing score frequency was significantly higher when the 3-level Angoff method was used compared to the standard Angoff method in these stations.

According to Table 6, results at the gavage station showed a higher frequency of pass scores in OSCE when using the 3-level Angoff method compared to the yes/no Angoff method. Conversely, at the IV therapy station, the frequency of pass scores was higher with the yes/no Angoff method than with the 3-level Angoff method, with a statistically significant difference.

Station and standard setting (score)	Pass, N (%)	Fail, N (%)	*P*-value
Pharmaceutical care			<0.001
Fixed score (12)	64 (99.8)	1 (0.2)	
Angoff (9.60)	65 (100.0)	0 (0.0)	
Vaccination			<0.001
Fixed score (12)	63 (96.9)	2 (3.1)	
Angoff (11)	65 (100.0)	0 (0.0)	
Gavage			<0.001
Fixed score (12)	52 (80.0)	13 (20.0)	
Angoff (11.60)	54 (83.1)	11 (16.9)	
The use of an ice bag			-
Fixed score (12)	61 (93.8)	4 (6.2)	
Angoff (10.80)	61 (93.8)	4 (6.2)	
IV therapy			< 0.001
Fixed score (12)	53 (81.5)	12 (18.5)	
Angoff (8.38)	62 (95.4)	3 (4.6)	
Venipuncture			-
Fixed score (12)	65 (100.0)	0 (0.0)	
Angoff (12.30)	65 (100.0)	0 (0.0)	
Personal protection			-
Fixed score (12)	65 (100.0)	0 (0.0)	
Angoff (13.20)	65 (100.0)	0 (0.0)	

Note: OSCE, objective structured clinical examination.

Table 1. Comparing the frequency of passing by fixed score and Angoff methods in nursing OSCE (N = 65).

Station and standard setting (score)	Pass, N (%)	Fail, N (%)	*P*-value
Pharmaceutical care			-
Fixed score (12)	64 (99.8)	1 (0.2)	
Yes/no Angoff (8.32)	65 (100.0)	0 (0.0)	
Vaccination			<0.001
Fixed score (12)	63 (96.9)	2 (3.1)	
Yes/no Angoff (13.80)	55 (84.6)	10 (15.4)	
Gavage			<0.001
Fixed score (12)	52 (80.0)	13 (20.0)	
Yes/no Angoff (12.30)	50 (76.9)	15 (23.1)	
The use of an ice bag			-
Fixed score (12)	61 (93.8)	4 (6.2)	
Yes/no Angoff (8.32)	65 (100.0)	0 (0.0)	
IV therapy			<0.001
Fixed score (12)	53 (81.5)	12 (18.5)	
Yes/no Angoff (6.15)	63 (96.9)	2 (3.1)	
Venipuncture			-
Fixed score (12)	65 (100.0)	0 (0.0)	
Yes/no Angoff (18.46)	7 (10.8)	58 (82.9)	
Personal protection			-
Fixed score (12)	65 (100.0)	0 (0.0)	
Yes/no Angoff (20)	12 (18.5)	53 (81.5)	

Note: OSCE, objective structured clinical examination.

Table 2. Comparing the frequency of passing by fixed score and yes/no Angoff methods in nursing OSCE (N = 65).

Station and standard setting (score)	Pass, N (%)	Fail, N (%)	*P*-value
Pharmaceutical care			-
Fixed score (12)	64 (99.8)	1 (0.2)	
3-level Angoff (10)	65 (100.0)	0 (0.0)	
Vaccination			-
Fixed score (12)	63 (96.9)	2 (3.1)	
3-level Angoff (11.50)	65 (100.0)	0 (0.0)	
Gavage			<0.001
Fixed score (12)	52 (80.0)	13 (20.0)	
3-level Angoff (10.76)	61 (93.8)	4 (6.2)	
The use of an ice bag			-
Fixed score (12)	61 (93.8)	4 (6.2)	
3-level Angoff (8.32)	65 (100.0)	0 (0.0)	
IV therapy			<0.001
Fixed score (12)	53 (81.5)	12 (18.5)	
3-level Angoff (5.38)	64 (98.5)	1 (1.5)	
Venipuncture			-
Fixed score (12)	65 (100.0)	0 (0.0)	
3-level Angoff (13.06)	65 (100.0)	0 (0.0)	
Personal protection			-
Fixed score (12)	65 (100.0)	0 (0.0)	
3-level Angoff (14.60)	65 (100.0)	0 (0.0)	

Note: OSCE, objective structured clinical examination.

Table 3. Comparing the frequency of passing by fixed score and 3-level Angoff methods in nursing OSCE (N = 65).

Station and standard setting (score)	Pass, N (%)	Fail, N (%)	*P*-value
Pharmaceutical care			-
Angoff (9.60)	65 (100.0)	0 (0.0)	
Yes/no Angoff (8.32)	65 (100.0)	0 (0.0)	
Vaccination			-
Angoff (11)	65 (100.0)	0 (0.0)	
Yes/no Angoff (13.80)	55 (84.6)	10 (15.4)	
Gavage			<0.001
Angoff (11.60)	54 (83.1)	11 (16.9)	
Yes/no Angoff (12.30)	50 (77.0)	15 (23.0)	
The use of an ice bag			-
Angoff (10.80)	61 (93.8)	4 (6.2)	
Yes/no Angoff (8.32)	65 (100.0)	0 (0.0)	
IV therapy			<0.001
Angoff (8.38)	62 (95.4)	3 (4.6)	
Yes/no Angoff (6.15)	63 (96.9)	2 (3.1)	
Venipuncture			-
Angoff (12.30)	65 (100.0)	0 (0.0)	
Yes/no Angoff (18.46)	7 (10.8)	58 (89.2)	
Personal protection			-
Angoff (13.20)	65 (100.0)	0 (0.0)	
Yes/no Angoff (20)	12 (18.5)	53 (81.5)	

Note: OSCE, objective structured clinical examination.

Table 4. Comparing the frequency of passing by Angoff and yes/no Angoff methods in nursing OSCE (N= 65).

Station and standard setting (score)	Pass, N (%)	Fail, N (%)	*P*-value
Pharmaceutical care			-
Angoff (9.60)	65 (100.0)	0 (0.0)	
3-level Angoff (10)	65 (100.0)	0 (0.0)	
Vaccination			-
Angoff (11)	65 (100.0)	0 (0.0)	
3-level Angoff (11.50)	65 (100.0)	0 (0.0)	
Gavage			<0.001
Angoff (11.60)	54 (83.1)	11 (16.9)	
3-level Angoff (10.76)	61 (93.8)	4 (6.2)	
The use of an ice bag			-
Angoff (10.80)	61 (93.8)	4 (6.2)	
3-level Angoff (8.32)	65 (100.0)	0 (0.0)	
IV therapy			<0.001
Angoff (8.38)	62 (95.4)	3 (4.6)	
3-level Angoff (5.38)	64 (98.5)	1 (1.5)	
Venipuncture			-
Angoff (12.30)	65 (100.0)	0 (0.0)	
3-level Angoff (13.06)	65 (100.0)	0 (0.0)	
Personal protection			-
Angoff (13.20)	65 (100.0)	0 (0.0)	
3-level Angoff (14.60)	65 (100.0)	0 (0.0)	

Note: OSCE, objective structured clinical examination.

Table 5. Comparing the frequency of passing by Angoff and the 3-level Angoff methods in nursing OSCE (N = 65).

Station and standard setting (score)	Pass, N (%)	Fail, N (%)	*P*-value
Pharmaceutical care			-
Yes/no Angoff (8.32)	65 (100.0)	0 (0.0)	
3-level Angoff (10)	65 (100.0)	0 (0.0)	
Vaccination			-
Yes/no Angoff (13.80)	55 (84.6)	10 (15.4)	
3-level Angoff (11.50)	65 (100.0)	0 (0.0)	
Gavage			<0.001
Yes/no Angoff (12.30)	50 (77.0)	15 (23.0)	
3-level Angoff (10.76)	61 (93.8)	4 (6.2)	
The use of an ice bag			-
Yes/no Angoff (8.32)	65 (100.0)	0 (0.0)	
3-level Angoff (8.32)	65 (100.0)	0 (0.0)	
IV therapy			<0.001
Yes/no Angoff (6.15)	63 (96.9)	2 (3.1)	
3-level Angoff (5.38)	61 (93.8)	4 (6.2)	
Venipuncture			-
Yes/no Angoff (18.46)	7 (10.8)	58 (89.2)	
3-level Angoff (13.06)	65 (100.0)	0 (0.0)	
Personal protection			-
Yes/no Angoff (20)	12 (18.5)	53 (81.5)	
3-level Angoff (14.60)	65 (100.0)	0 (0.0)	

Note: OSCE, objective structured clinical examination.

Table 6. Comparing the frequency of passing by yes/no Angoff and the 3-level Angoff methods in nursing OSCE (N = 65).

4. Discussion

This study was conducted to compare the frequency of passing of nursing students participating in OSCE by different standard-setting methods (fixed score, Angoff, yes/no Angoff, and the 3-level Angoff). Different studies have been undertaken to compare different standard settings for the objective structured clinical exam.[10,11,19]

As findings revealed, there was a significantly higher rate of passing scores in OSCE for pharmaceutical care, vaccination, gavage, and IV therapy stations when the Angoff method was used to compare the fixed score method. No significant difference was observed in other stations. This may be attributed to the lower standard passing score achieved through the Angoff method compared to the fixed score method in these specific stations. Additionally, the fixed score method relies on a predetermined standard for the test without considering individual examinee difficulty levels. This fixed standard may lead to varying passing rates among students in different exams, as it does not adjust based on student performance. It should be noted that, in this study, the Angoff score was lower than the fixed score in the station using an ice bag; however, the rate of successful completion was the same.

The study undertaken by Jalili and Mortaz Hejri to determine the quorum score and acceptance rate of pre-internship medical students at Tehran University of Medical Sciences in OSCE using 4 methods of fixed score, Angoff, borderline regression, and Cohen showed that despite the acceptance standard of Angoff's method in 2 stations of genital examination and suturing, which was lower than the borderline regression method, the frequency of acceptance in these stations was the same.[10]

Findings also showed that there is a significant difference in passing rates on OSCE assessments using the fixed score method at vaccination, gavage, and IV therapy stations compared to the yes/no Angoff method. Specifically, the passing rates were higher using the fixed score method at the vaccination and gavage stations, while the yes/no Angoff method resulted in higher passing rates at the IV therapy station. We attribute the results obtained for vaccination and gavage stations to the lower passing score of the fixed score method. Similarly, the same result obtained for the IV therapy station can be due to the lower passing score obtained from the yes/no Angoff method. We assert that judges for the yes/no Angoff method take the difficulty or ease of the items into account and, based on the borderline student performance, allocate a score from 0 to 100. Consequently, it seems that judges found the performance at the IV therapy station to be more challenging compared to the vaccination and gavage stations.

Another finding of the present study is that the rate of achieving a passing score in the nursing OSCE, particularly in the gavage and IV therapy stations, is significantly higher when using the 3-level Angoff method compared to the fixed score method. This finding may be attributed to the lower standard-setting score and the perceived simplicity of the 3-level Angoff method in comparison to the fixed score method.

Another finding from this research indicated a significant increase in acceptance rates during OSCE evaluations at the gavage station when using the Angoff score method as opposed to the yes/no Angoff method. In contrast, a higher frequency of passing scores was observed in the IV therapy station when using the yes/no Angoff method compared to the Angoff method. This difference was found to be statistically significant. This discrepancy may be attributed to the lower acceptance standard score in the Angoff method at the gavage station and the leniency of judges in the yes/no Angoff method at the IV therapy station.

Another finding of this study showed a significant difference in the frequency of passing scores in OSCE at gavage and IV therapy stations using the Angoff method compared to passing scores obtained through the 3-level Angoff method. In the gavage and IV therapy stations, the frequency of passing scores using the 3-level Angoff score method was significantly higher than the passing scores using the Angoff scoring method. One

possible explanation for this phenomenon is the lower standard scores obtained by the 3-level Angoff method compared to the Angoff method. It is hypothesized that the judges using the 3-level Angoff method are limited to selecting scores of 0%, 50%, or 100% for borderline students at a station. If they do not choose 100, they must choose either 0% or 50%. However, judges using the Angoff method have the flexibility to choose scores within a range of 0–100. As a result, there is a likelihood that judges using the Angoff method may select scores between 50 and 100, leading to a higher passing score. Kaufman et al.[24] believe that Angoff's method may lead to setting a very high and strict standard.

Our research findings indicate that there was a significantly higher frequency of passing scores in the OSCE at the gavage station when using the 3-level Angoff method compared to the Yes/No Angoff method. Conversely, at the IV therapy station, the frequency of passing scores using the Yes/No Angoff method was higher than those using the 3-level Angoff method, with statistical significance. We suggest that the discrepancy in passing rates between these 2 stations is a result of the lower acceptance standard set by the 3-level Angoff method at the gavage station and a lower acceptance standard score at the IV therapy station.

Limitations

This study has its own limitations. They include: (1) a lack of experienced panelists who are not familiar with the standard setting for the nursing OSCE. (2) focus on determining the acceptance standards based on test-oriented approaches rather than examinee-oriented methods. (3) the paucity of relevant studies in line with our research. (4) Limited external validity of study results. While the findings on the comparative effectiveness of different standard-setting methods in the OSCE are interesting, they may not apply universally to other nursing programs, institutions, or countries. The sample size is small, and the sample itself may not be fully representative of the broader population of nursing students and educators. Additionally, the specific clinical stations and methods tested may not be applicable to other contexts with different priorities or curricula. Further studies with larger and more diverse samples, conducted across multiple settings, would be necessary to better understand the generalizability of these findings.

Future research can focus on the design and implementation of studies to determine the relationship between the pass/fail scores and their frequency, using different methods such as borderline regression, Cohen, etc.

5. Conclusions

This study examined 4 different approaches to standard setting: fixed score, Angoff, yes/no Angoff, and 3-level Angoff, to analyze the passing rates associated with each method. Due to the diversity in standard-setting approaches and OSCE stations' design, it was not possible to find similar studies that investigated the acceptance rates across different stations. Nonetheless, it is evident that despite judges' consistency in scoring methods, varying approaches in determining acceptance standards and judges' strictness or leniency can result in disparate scores and acceptance frequencies. These discrepancies, which have been noted in previous studies, can be statistically significant.

Acknowledgments

We would like to thank all the nursing teachers and students who participated in this study.

Author's contributions

All authors participated in this study; MY is the principal investigator and the main author who supervised the process of study. FM, ShP, KM, and AB collected data. MY and AB analyzed the data and wrote the final draft of the manuscript. FM, ShP, and KM revised the manuscript and finally approved the manuscript before submission. The authors read and approved the final manuscript.

Data availability

The data of this study are available upon request to the corresponding author.

Ethics approval

This study was financially supported by the vice-chancellery for research of Bushehr University and the National Agency for Strategic Research in Medical Sciences Education (NASR) of Tehran (Code: IR.NASRME.REC.1402.021). It is important to mention that this study was based on a master's thesis in nursing, which was approved by the ethics committee of Bushehr University of Medical Sciences (IRB approval number: IR.BPUMS.REC.1400.086).

Conflicts of interest

All contributing authors declare no conflicts of interest.

References

1. French S, Dickerson A, Mulder RA. A review of the benefits and drawbacks of high-stakes final examinations in higher education. *High Educ*. 2024;88:893–918.
2. Boud D, Falchikov N. Aligning assessment with long-term learning. *Assess Eval High Educ*. 2006;31:399–413.
3. Prakash Jyoti, Chatterjee K, Srivastava K, et al. Workplace based assessment: a review of available tools and their relevance. *Industrial Psychiatry Journal*. 2020;29:200–204.
4. Pront L, McNeill L. Nursing students' perceptions of a clinical learning assessment activity: 'linking the puzzle pieces of theory to practice'. *Nurse Educ Pract*. 2019;36:85–90.
5. Brannick MT, Erol-Korkmaz HT, Prewett M. A systematic review of the reliability of objective structured clinical examination scores. *Med Educ*. 2011;45:1181–1189.
6. Maghsoodi F, YazdanFard M, Pouladi S, et al. Using the methods of fixed score, Angoff, yes/no Angoff and three-level Angoff to determine the standard and acceptance rate of nursing students participating in an objective structured clinical exam. *Iran J Med Educ*. 2023;23:217–224 (in Persian).
7. Dewan P, Khalil S, Gupta P. Objective structured clinical examination for teaching and assessment: evidence-based critique. *Clin Epidemiol Glob Health*. 2024;25:101477.
8. Chabrera C, Diago E, Curell L. Development, validity and reliability of objective structured clinical examination in nursing students. *SAGE Open Nurs*. 2023;9:1–8.
9. Cowan DT, Norman I, Coopamah VP. Competence in nursing practice: a controversial concept – A focused review of literature. *Nurse Educ Today*. 2005;25:355–362.
10. Jalili M, Mortazhejri S. Standard setting for objective structured clinical exam using four methods: prefixed score Angoff borderline regression and cohens. *Strides Dev Med Educ*. 2012;9:77–84 (in Persian).
11. Boursicot KA, Roberts TE, Pell G. Using borderline methods to compare passing standards for OSCEs at graduation across three medical schools. *Med Educ*. 2007;41:1024–1031.
12. Kilminster S, Roberts T. Standard setting for OSCEs: trial of borderline approach. *Adv Health Sci Educ Theory Pract*. 2004;9:201–209.
13. Reid KJ, Dodds A. Comparing the borderline group and borderline regression approaches to setting Objective Structured Clinical Examination cut scores. *J Contemp Med Edu*. 2014;2:8–12.
14. Schoonheim-Klein M, Muijtjens A, Habets L, et al. Who will pass the dental OSCE? Comparison of the Angoff and the borderline regression standard setting methods. *Eur J Dent Educ*. 2009;13:162–171.
15. Cizek GJ, Bunch MB. Standard Setting: A Guide to Establishing and Evaluating Performance Standards for Tests. Thousand Oaks, CA: Sage Publications; 2007:5–11.
16. Ho TK, Abu Kassim NL, O'Malley L, et al. Standard setting for dental knowledge tests: reproducibility of the modified Angoff and Ebel method across judges. *BMC Med Educ*. 2025;25:1–13.
17. Dwyer T, Wright S, Kulasegaram KM, et al. How to set the bar in competency-based medical education: standard setting after an Objective Structured Clinical Examination (OSCE). *BMC Med Educ*. 2016;16:1.
18. Kramer A, Muijtjens A, Jansen K, et al. Comparison of a rational and an empirical standard setting procedure for an OSCE. *Med Educ*. 2003;37: 132–139.
19. Jalili M, Hejri SM, Norcini JJ. Comparison of two methods of standard setting: the performance of the three-level Angoff method. *Med Educ*. 2011;45:1199–1208.
20. Park J. Possibility of using the yes/no Angoff method as a substitute for the percent Angoff method for estimating the cut score of the Korean Medical Licensing Examination: a simulation study. *J Educ Eval Health Prof*. 2022;19:1–9.
21. Hejri SM, Jalili M, Muijtjens AM, et al. Assessing the reliability of the borderline regression method as a standard setting procedure for objective structured clinical examination. *J Res Med Sci*. 2013;18:887–891.
22. Dwivedi NR, Vijayashankar NP, Hansda M, et al. Comparing standard setting methods for objective structured clinical examinations in a caribbean medical school. *J Med Educ Curric Dev*. 2020;7: 1–10.
23. Yim MK, Shin S. Using the Angoff method to set a standard on mock exams for the Korean Nursing Licensing Examination. *J Educ Eval Health Prof*. 2020;17:14.
24. Kaufman DM, Mann KV, Muijtjens AM, et al. A comparison of standard-setting procedures for an OSCE in undergraduate medical education. *Acad Med*. 2000;75:267–271.