ASIAN JOURNAL OF EDUCATIONAL SCIENCES
ISSN: 0866-7675
eISSN: 2815-567X
DONG THAP UNIVERSITY JOURNAL OF SCIENCE - DONG THAP UNIVERSITY
https://ajesdthu.vn

DOI: https://doi.org/10.52714/dthu.ajes.30.2052

Applying the Nedelsky Method to Determine Passing Scores for End-of-Course Examinations in Dentistry

Hung Manh Vo, Truc Thi Hoang Nguyen, Nhi Thi Phuong Pham, Lien Thi Nguyen and Vu Hoa Anh Dien*

Faculty of Dentistry, Van Lang University, Vietnam

**Corresponding author, Email: vu.dha@vlu.edu.vn*

Article Info.

Received: August 10, 2026
Revised: September 19, 2026
Accepted: October 16, 2026

Keywords

competency-based assessment, dental education, Nedelsky method, passing score, standard setting

Cite

Vo, H. M., Nguyen , T. T. H., Pham, N. T. P., Nguyen , L. T., & Vu, H. A. D. (2026). Applying the Nedelsky Method to Determine Passing Scores for End-of-Course Examinations in Dentistry. *Dong Thap University Journal Science - Asian Journal of Educational Sciences, 1*(3), 15-27. https://doi.org/10.52714/dthu.ajes.30.2052

Abstract

This research establishes a scientifically defensible passing score for the Oral and Maxillofacial Pathology final examination for dental students at Van Lang University by using the Nedelsky method. Five experts independently reviewed a 50-item multiple-choice test to identify distractors that a minimally competent candidate (MCC) could reasonably eliminate. The Minimum Passing Level (MPL) was calculated by averaging these expert estimates and subsequently applied to the examination results of 132 students. Student scores followed a normal distribution ($p = 0.525$), with an overall mean of 5.4 ± 1.3. The panel established an MPL of 52.4%, requiring at least 27 correct answers to pass. Compared to arbitrary fixed thresholds of 40% and 50%, the student pass rate decreased from 87.9% and 67.4%, respectively, to 52.3%. Twenty students - 15.2% of the cohort and 22.5% of those passing under the 50% threshold - were reclassified as borderline, indicating that they did not meet the competence level expected by the expert panel. The Nedelsky method created a higher passing standard based on student competence rather than convention. These results support adopting criterion-referenced methods to ensure assessments are fair and consistent with competency-based dental education.

Copyright © 2026 The author(s). This work is licensed under a CC BY-NC 4.0 License

1. INTRODUCTION

In the global landscape of health professions education, particularly in dentistry, the shift toward competency-based curricula has necessitated a more rigorous approach to assessment. The fundamental goal is to ensure that graduates possess the minimum level of competence required for safe and ethical clinical practice. Consequently, establishing a passing score for end-of-course examinations is a high-stakes decision that directly impacts public safety and professional standards. However, a significant challenge remains: many institutions continue to rely on arbitrary fixed passing scores (e.g., 50% or 60%). These traditional thresholds often lack a solid theoretical foundation and fail to accurately reflect the actual difficulty of the exam or the specific competencies required (Downing & Yudkowsky, 2009; Norcini, 2003).

Competency-based dental education has been increasingly adopted by dental schools worldwide as a framework that shifts the focus from time-served curricula to the demonstration of defined professional competencies (Albino et al., 2008). Within this framework, summative assessments, particularly end-of-course examinations, serve as critical mechanisms that determine whether a student has achieved the minimum proficiency required for safe patient care. Furthermore, evidence from dental schools in Malaysia has shown that most institutions continue to use an arbitrary fixed passing mark of 50% for their final professional examinations without structured standard-setting procedures, highlighting that this problem is widespread across the region (Ho et al., 2025a, 2025b). This situation underscores that the adoption of criterion-referenced methods is not merely a best practice but an essential step toward defensible, competency-aligned assessment.

In Vietnam, the adoption of criterion-referenced standard-setting methods is still in its nascent stages. While international evidence suggests that the choice of method can significantly alter pass-fail outcomes, local data remain sparse. At Van Lang University, Oral and Maxillofacial Pathology 1 serves as a critical foundation for clinical diagnosis; yet, its passing standards have historically been based on convention rather than empirical evidence. This research gap creates an urgent need for a defensible, evidence-based threshold. To specify what remains unknown, we are aware of no published study that has applied the Nedelsky method to an Oral and Maxillofacial Pathology examination within a Vietnamese dental programme. Three questions therefore remain unanswered: (1) What passing score does the Nedelsky procedure yield for a 50-item MCQ examination in this discipline? (2) How does this passing score compare with the 40% and 50% thresholds currently applied by institutional regulation and convention, respectively? (3) How many candidates would change their pass/fail status when the passing standard is derived from expert judgement rather than convention?

Within Southeast Asia, the transition toward outcome-based and competency-driven dental education has been uneven. While countries such as Sweden and Malaysia have begun integrating structured standard-setting procedures into licensing and professional qualifying examinations, Vietnamese dental schools have largely maintained traditional, convention-based approaches to score determination (Abd-Rahman et al., 2021; Dalum et al., 2023). In Malaysia, most dental institutions continue to arbitrarily apply a fixed 50% passing threshold for final professional examinations, a practice that faculty members themselves have identified as problematic for ensuring graduate competence (Ho et al., 2025b). The Ministry of Education and Training (MOET) of Vietnam currently permits individual institutions to set their own passing thresholds, often defaulting to fixed percentages without explicit reference to minimum competency benchmarks. This institutional autonomy, while flexible, introduces heterogeneity in assessment rigor across programs and makes it difficult to ensure a consistent

standard of graduate readiness for clinical practice. In this context, discipline-specific studies documenting the empirical impact of criterion-referenced methods are critically needed to build a local evidence base and to inform future policy reform in Vietnamese health professions education.

Therefore, this study aims to: (1) apply the Nedelsky method to establish a criterion-referenced passing score for third-year dental students at Van Lang University, and (2) evaluate the theoretical pass rates resulting from this standard. By doing so, this study contributes to the development of fairer, more consistent assessment practices in Vietnamese dental education.

2. LITERATURE REVIEW

Educational measurement distinguishes between two fundamental assessment philosophies: norm-referenced and criterion-referenced approaches. Norm-referenced standards are relative, defining success by comparing a student's performance to that of their peers, such as selecting the top 20% or using a "mean minus one standard deviation" formula. While useful for competitive selection, this approach is considered inadequate for certifying professional competence because it does not guarantee a fixed level of mastery and can be influenced by the overall ability of a specific cohort. Conversely, criterion-referenced standards (absolute standards) measure performance against a pre-defined level of competence regardless of group performance. In health professions education, where verifying absolute competency is essential for clinical safety, criterion-referenced standard-setting methods are widely recommended for competence-certifying examinations (Downing et al., 2006; Norcini, 2003), although no single standard-setting procedure is universally accepted as definitive.

First introduced by Leo Nedelsky in 1954, Nedelsky method (NM) remains a prominent absolute standard-setting procedure specifically designed for multiple-choice examinations. Its unique theoretical premise focuses on distractor elimination. Subject matter experts (SMEs) are required to review every item and identify which incorrect alternatives (distractors) a minimally competent candidate (MCC) - defined as a student who has just enough knowledge to pass, could reasonably recognize as false. The probability of an MCC choosing the correct answer is then estimated as the reciprocal of the remaining plausible options. This granular, item-level analysis provides a detailed cognitive model of student performance, grounding the final passing score in expert judgment rather than arbitrary convention (Ho et al., 2025a).

The reliability of any absolute standard depends heavily on the expertise and training of the judging panel. Recommended panel sizes typically range from 5 to 30 members, with larger groups generally improving the reproducibility of the passing score (Downing et al., 2006). A recurring theme in contemporary research is the "Failure to Fail" phenomenon, where arbitrary fixed thresholds (e.g., 40%, 50%, or 60%) allow incompetent candidates to pass simply because the exam was too easy, or conversely, fail competent students if the exam was unexpectedly difficult. Studies in dental education have challenged these arbitrary fixed scores, demonstrating that evidence-based methods often yield thresholds that differ significantly from institutional conventions, sometimes impacting pass rates by as much as 41% (Yousef et al., 2020).

Central to NM is the construct of MCC. This is a theoretical construct representing the hypothetical examinee who possesses precisely the minimum level of knowledge, skill, and judgment deemed necessary to practice safely and effectively (Mortaz & Jalili, 2014). Critically, MCC is not an average student, nor the weakest student, but rather a carefully defined reference point that anchors the entire standard-setting exercise. Experts are asked not "what score should students achieve?" but rather "which distractors could a student with just-sufficient knowledge plausibly eliminate?" This cognitive reframing makes the judgment task

more concrete and reduces the tendency for raters to be influenced by the actual distribution of observed scores (Abd-Rahman et al., 2021). Achieving consensus on the MCC profile prior to item review is therefore considered a best-practice step that improves both the validity and reproducibility of Nedelsky-based passing scores (Kim & Yang, 2020).

Among the most frequently compared criterion-referenced methods in health professions education are NM, the Angoff method, and the Ebel method. The modified Angoff method requires panelists to estimate, for each item, the probability (expressed as a proportion from 0 to 1) that a MCC would answer correctly, then averages these estimates across items and panelists to produce a passing score. In contrast, the Ebel method requires panelists to first classify items along two dimensions, relevance and difficulty, and then estimate the expected MCC performance within each cell of the resulting matrix. A recent systematic review and simulation study found that the choice of standard-setting method can produce passing scores differing by up to 15% points on the same examination, with NM tending to yield slightly lower thresholds than Angoff on most MCQ-based assessments (Ho et al., 2025a). This variability underscores the importance of method selection, as well as panel training, in ensuring that passing cores are consistent and defensible across examination cycles. For institutions introducing structured standard setting for the first time, NM offers a practical starting point due to its strong item-level grounding and relatively straightforward implementation process, particularly when the examination format is exclusively multiple-choice (Yousef et al., 2020).

Evidence from dental education contexts outside Vietnam is still growing but consistently points to the same conclusion: structured standard-setting methods identify a non-trivial proportion of students who pass under conventional thresholds yet fall below the competency level defined by subject matter experts. In Malaysia, a study comparing multiple standard-setting methods for a professional qualifying dental examination found that evidence-based passing scores differed substantially from the arbitrary 50% institutional threshold, with important implications for graduate readiness (Abd-Rahman et al., 2021). Similarly, research from Sweden demonstrated that applying a structured Angoff-based standard-setting procedure to theoretical proficiency tests for dentists was feasible and defensible, effectively differentiating candidates who met licensure requirements from those who did not (Dalum et al., 2023). These findings from different geographic and educational contexts reinforce the urgency of adopting criterion-referenced methods in Vietnamese dental education and provide a comparative benchmark against which the outcomes of the present study can be interpreted.

In Vietnam, the application of criterion-referenced methods in medical and dental education is still in its nascent stages. Most institutions continue to rely on arbitrary fixed passing scores that lack a clear theoretical basis and may not accurately reflect the actual difficulty of specialized courses. While recent domestic studies have begun exploring NM in general medicine, there is a significant lack of data regarding its performance within the specific multidisciplinary context of Oral and Maxillofacial Pathology. This study addresses these gaps by establishing a defensible, scientifically-based passing score for dental students in Vietnam. By synthesizing expert judgment with local student performance data, this research contributes to a fairer and more reliable assessment framework that ensures only truly competent candidates progress to clinical practice.

3. MATERIALS AND METHODS

This study was conducted in two stages with two different study populations.

Stage 1: Determining the passing score:

The study population in this stage was the full set of multiple-choice questions (MCQs) used in the Oral and Maxillofacial Pathology 1 examination.

Stage 2: Determining the pass rate:

The study population in this stage included third-year dental students at Van Lang University who took the theoretical examination for the Oral and Maxillofacial Pathology 1 course on May 8, 2025.

Study Design: This study used a descriptive cross-sectional design. This was a two-stage methodological standard-setting study. Stage 1 comprised a prospective expert-judgement exercise applied to all 50 examination items, while Stage 2 involved a retrospective analysis of the anonymised student scores from the same examination, in which the passing score derived in Stage 1 was applied post hoc alongside the two conventional thresholds.

Sampling Method: The sampling method was total population sampling, including all third-year Dental students of the 2021-2027 cohort at Van Lang University who took the final theoretical examination of the Oral and Maxillofacial Pathology 1 course on May 8, 2025 (n = 132). The 40% threshold used for comparison is the minimum passing score stipulated in the credit-based academic regulations of Van Lang University, under which a score of 4.0 on a 10-point scale is the minimum required for course credit; the 50% threshold is the conventional mark applied in practice.

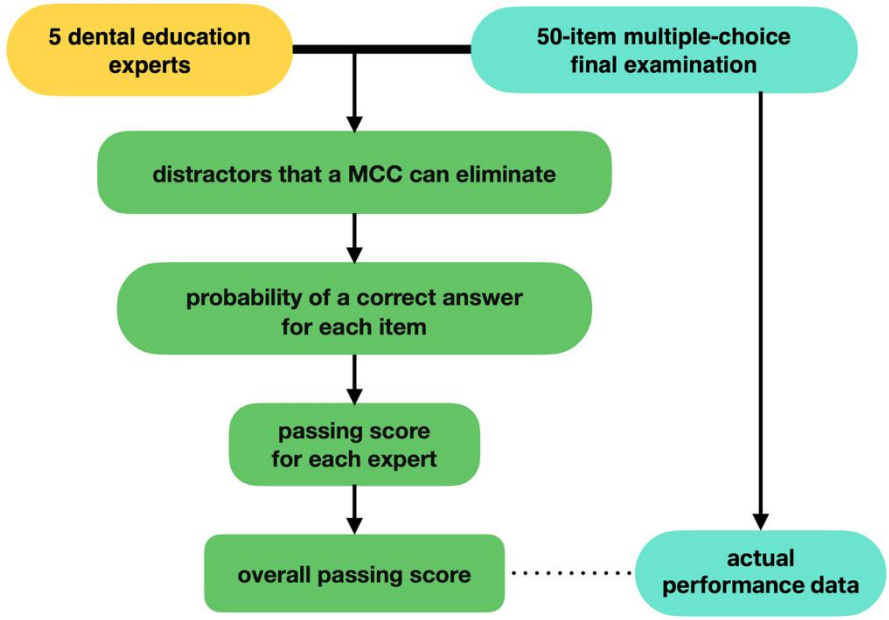


Figure 1. Flowchart of the Nedelsky Standard-Setting Procedure

Procedure:

Step 1: Defining Competencies: In this study, we defined the competency to be assessed as "Basic knowledge of the Oral and Maxillofacial Pathology 1 course for third-year dental students". The assessment format was a multiple-choice examination (MCQ), using Test Version 1 administered on May 8, 2025. Before reviewing the items, the panel agreed on a written description of MCC, which remained visible throughout the rating session.

Step 2: Establishing the Expert Panel: An expert panel of five faculty members from the Department of Dentistry participated in the standard-setting process using NM. Panelists were selected by the Head of the Department based on their academic expertise and teaching experience. All panel members held postgraduate qualifications and had formal training in health professions education. They were also familiar with the Nedelsky standard-setting approach and had extensive experience teaching the Oral and Maxillofacial Pathology course, which enabled them to judge the expected performance of MCC.

Step 3: Nedelsky standard-setting process: The standard-setting procedure followed several steps.

First, all panelists received structured training on the Nedelsky method, including an overview of the methodology and practice with sample questions.

Second, each panelist independently reviewed all MCQ items and determined which distractors a minimally competent student would likely be able to eliminate. Panellists worked independently and were blinded to one another's judgements at this stage. For each item, the probability of a correct response for an MCC was calculated as the reciprocal of the number of plausible options remaining after eliminating the implausible distractors.

Third, panelists met to discuss their judgments and reach agreement on the number of remaining plausible options for each item.

For each question, the probability of a correct response for an MCC was estimated based on the number of remaining options. The final passing score for the examination was calculated as the average of the Nedelsky values across all items.

Data Analysis: Data were entered and analyzed using STATA version 17.0. Both descriptive and analytical statistics were used. The passing scores determined by the expert panel were analyzed to calculate the mean, standard deviation, and the minimum and maximum values. The pass and fail rates based on the Nedelsky passing score were also calculated and compared with previous pass rates from the same examination.

Ethical Considerations: The study involved no intervention and posed no risk to participants. The research protocol was reviewed and approved by the Research ethics Committee of Van Lang University (No. 27/2026/IRB-VN01.078; Date of appoval: March 18th, 2026). All student data were anonymized prior to analysis and were used only for research purposes. Confidentiality was strictly maintained throughout the study in accordance with ethical principles for research involving human participants.

4. RESULTS

4.1. Distribution of Examination Scores

A total of 132 students completed the assessment. The mean score was 5.4 ± 1.3 on a 10-point scale, indicating a moderate overall performance. The Shapiro-Wilk test demonstrated that the score distribution did not depart significantly from normality ($p = 0.525$). Parametric tests were subsequently applied. Visual inspection of the histogram confirmed a symmetric bell-shaped curve. This distributional pattern describes the observed spread of scores only. Since item-level psychometric indices were not available for this examination, no inferences about item discrimination or test difficulty were drawn from the shape of the score distribution.

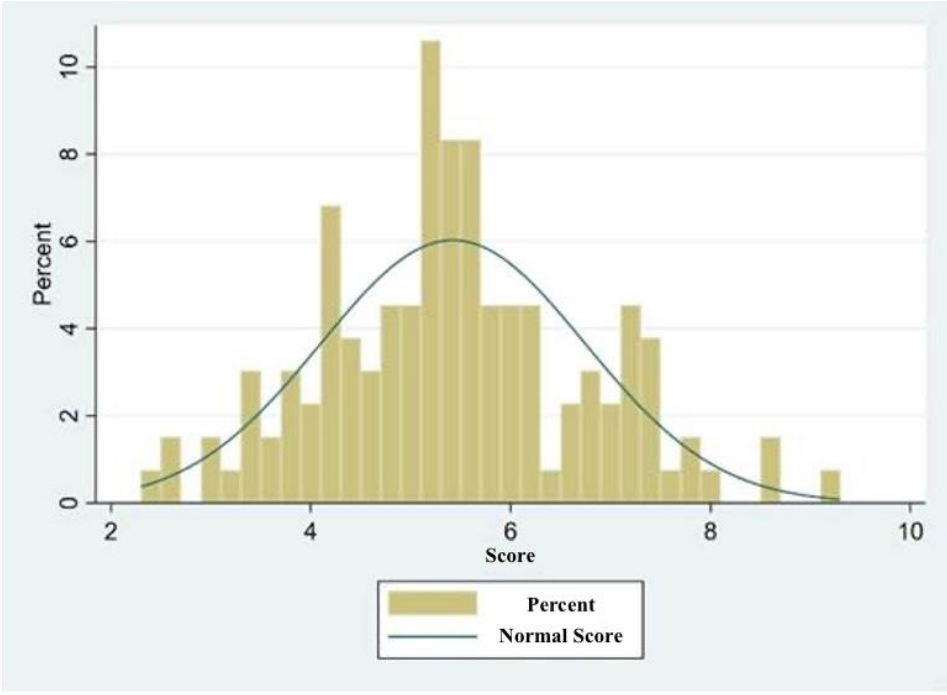


Figure 2. Distribution of Students' Scores

4.2. Minimum Passing Level Determination Using the Nedelsky Method

Five expert panelists independently evaluated all 50 multiple-choice items using NM procedure. The recommended minimum passing thresholds varied from 38% to 69%, reflecting differences in expert judgment regarding the likelihood that a borderline candidate could eliminate distractors. The aggregated Minimum Passing Level (MPL) was 52.4%, corresponding to 26.19 correct responses. Using conventional rounding procedures to maintain analytical precision, the required passing score for passing was set at 27 out of 50 items. The conversion of the MPL into an operational passing score was as follows: 52.38% × 50 = 26.19 items. Since the number of correct responses must be an integer, the operational passing score was set at the smallest integer greater than or equal to 26.19, namely 27 correct responses. Thus, a candidate answering 26 items correctly (52.0%) falls below the panel-derived MPL and does not pass. The operational passing score of 27 out of 50 corresponds to 54.0% of the maximum score, which is 1.6 percentage points above the panel-derived MPL of 52.4%.

Table 1. Results of Five Expert Panelists Independently Evaluated All Items Using the Nedelsky Procedure

Expert	Mean	Standard deviation
Expert 1	0.69	0.29
Expert 2	0.40	0.12
Expert 3	0.66	0.26
Expert 4	0.50	0.25
Expert 5	0.38	0.08

4.3. Comparison of Pass/Fail Rates Across Standard-Setting Approaches

Student outcomes were compared under three different standard-setting methods: the institutional credit-based regulation (40%), the traditional fixed passing score (50%), and NM (52.38%).

Under the credit-based regulation, 87.9% of students achieved the threshold of 20 correct items, resulting in a failure rate of only 12.1%. Applying the traditional 50% passing score reduced the pass rate to 67.4% (89 students). In contrast, the Nedelsky standard set at 27 correct items yielded a markedly lower pass rate of 52.3%, with 63 students failing the examination.

A distinct borderline group of 20 students scored 25–26 items, meeting the traditional 50% criterion but falling below the Nedelsky standard. These students represented 15.2% of the whole cohort (20/132) and 22.5% of the 89 students who would have passed under the 50% threshold (20/89). Under the Nedelsky standard, these candidates were therefore classified as not meeting the passing standard for this examination, indicating that their observed performance on this assessment fell below the level defined by the expert panel. This reclassification refers specifically to performance on this single MCQ examination and should not be interpreted as a broader judgement of overall competence.

Table 2. Comparison of Student Outcomes under Three Standard-Setting Approaches

Method	Passing threshold (%)	Correct items *(out of 50)*	No. Passed	Pass rate (%)	No. Failed	Fail rate (%)
Credit-based regulation	40.0%	20	116	87.9%	16	12.1%
Traditional convention	50.0%	25	89	67.4%	43	32.6%
Nedelsky method	52.4%	27	69	52.3%	63	47.7%

5. DISCUSSION

5.1. Impact of the Nedelsky Method

In health professions education, especially in dental education, setting an appropriate passing score is an important part of assessment because it affects decisions about whether students have achieved the minimum competence needed for safe clinical practice (Downing et al., 2006). Standard-setting methods such as NM can help reduce subjectivity in this process. In many examinations, passing scores are still set at round values such as 50% or 60%, without considering item difficulty or the expected competence level of students (Ho et al., 2025). These practices may not fully reflect the minimum level of knowledge required in the discipline. Therefore, criterion-referenced methods such as the Nedelsky approach have been recommended to establish clearer and defensible passing standards (Yousef et al., 2017).

In the present study, NM produced a passing score of 52.38%, slightly higher than the traditional 50% rule. This corresponds to answering 27 out of 50 questions correctly, rather than 25 as required by the conventional rule. Applying this standard identified a borderline group of 20 students - 15.2% of the cohort and 22.5% of those who would have passed under

the traditional 50% rule - whose performance fell below the passing standard defined by the expert panell. Meeting or not meeting this passing score reflects the minimum knowledge expected on this written examination and should not be interpreted as a measure of overall clinical competence. This result suggests that fixed passing scores may not always represent the expected knowledge level of minimally competent candidates. By evaluating the plausibility of distractors in each question, NM provides a more structured way to estimate expected student performance (Downing et al., 2006; Ho et al., 2025).

The identification of this borderline group also changed in the examination results. When comparing the traditional 50% rule with the Nedelsky standard, the number of failing students increased from 43 to 63. This result indicates that a group of students was performing close to the minimum competence level. In medical and dental education, where examination outcomes often influence progression decisions, structured standard-setting methods can help identify borderline students more carefully. In contrast, lower thresholds such as the 40% credit-system regulation or the conventional 50% rule may allow some students with insufficient knowledge to pass. This situation has sometimes been described as the "failure to fail" phenomenon.

When viewed within the broader context of standard-setting methods, NM has both advantages and limitations. Compared with norm-referenced methods such as the Cohen method, the Nedelsky approach is more consistent with competency-based education because the passing score does not depend on the performance of top students. Norm-referenced methods may cause the passing standard to change from one cohort to another, which may not reflect the minimum competence required for clinical practice.

Among criterion-referenced methods, the Angoff method is currently the most widely used approach in medical education because it can be applied to many types of assessment formats. Previous studies have often reported that NM tends to produce lower passing scores than Angoff. This difference is related to the mechanisms of the two approaches. The Nedelsky method focuses on eliminating distractors, whereas Angoff requires experts to estimate the probability that a minimally competent candidate (MCC) will answer the item correctly (Chang, 1999). However, in our study NM resulted in a slightly higher threshold than the traditional standard. This may reflect careful evaluation of item structure by the expert panel, which led to a higher expectation of student knowledge.

Applying the Nedelsky method also improves transparency because each exam item is evaluated individually. This detailed documentation can provide useful evidence when responding to academic appeals or accreditation requirements. Overall, the use of criterion-referenced methods instead of fixed passing scores can support more reliable and defensible assessment practices in competency-based dental education (Norcini, 2003; Verhoeven et al., 2002).

5.2. Limitations of the Nedelsky Method

Despite these advantages, NM has several limitations. One commonly discussed issue in educational measurement is that the method may produce lower passing scores compared with other criterion-referenced methods such as Angoff (Chang, 1999). In high-stakes examinations, where verifying a high level of competence is essential, this potential tendency toward lower thresholds should be interpreted carefully. Although the passing score in our study (52.4%) was higher than the traditional 50% rule, previous research suggests that Nedelsky-based standards may sometimes differ from those obtained using other approaches in professional certification (Nedelsky, 1954; Verhoeven et al., 2002).

Another limitation is that NM is mainly suitable to MCQ because it relies on the evaluation of distractors to estimate the probability of a correct response. In modern health professions education, assessment often includes various formats such as essays, clinical performance tasks, or Objective Structured Clinical Examinations (OSCEs). In these contexts, methods such as Angoff may be more flexible (Ho et al., 2025a).

The method also assumes that students eliminate incorrect options and then guess among the remaining answers. In practice, student behavior during examinations may be more complex. For example, students may have partial knowledge or may eliminate the correct answer because of misunderstanding. These situations cannot be fully captured by the probabilistic structure of NM (Yousef et al., 2020).

In our study, the identification of a 15% borderline group shows that the method is sensitive to item-level difficulty. However, because 47.7% of students failed under the Nedelsky standard, any inaccuracies in the assumptions about guessing behavior could influence pass-fail decisions. Therefore, although the method provides a structured and transparent approach to standard setting, combining it with other methods may be useful in high-stakes assessments (Chang, 1999; Yousef et al., 2020).

5.3. The Hybrid Approach (Angoff-Nedelsky Combination)

Because no single procedure is universally accepted, some authors have proposed hybrid approaches such as the Angoff–Nedelsky combination, which tends to produce a passing score between those generated by the two individual methods (Chang, 1999). Evaluating such a hybrid approach in the present examination would be a useful direction for future research but was beyond the scope of the current study.

Due to standard setting involves expert judgment, variation among panelists is expected. In our study, the proposed passing scores ranged from 38% to 69%, reflecting differences in expert opinions. To reduce potential classification errors, some institutions apply statistical adjustments such as the Standard Error of Measurement (SEM) to the final passing score (Kim & Yang, 2020). These adjustments can help improve the reliability of the passing standard. In dental education, where confirming minimum competence is closely related to patient safety, such approaches may support more consistent and appropriate assessment decisions (Downing & Yudkowsky, 2009).

5.4. Strengths and Limitations of the Study

One strength of this study is the expertise of the judging panel. The panel consisted of five experienced faculty members with postgraduate qualifications and formal training in health professions education. Their teaching experience in the Oral and Maxillofacial Pathology 1 course allowed them to make informed judgments about the expected performance of minimally competent candidates. In addition, the standard-setting process followed a structured procedure that included panel training, independent item review, and consensus discussion, which helped improve the objectivity of the results.

However, several limitations should also be acknowledged. *First,* the expert panel consisted of only five members, which is at the lower end of the recommended range for NM (George et al., 2006; Shulruf et al., 2016). A larger and more diverse panel may have provided additional perspectives and further refined the final passing score (Yousef et al., 2020; Bahammam, 2017). *Second,* the analysis focused on a single course examination using four-option MCQs, which may limit the generalizability of the findings to other assessment formats used in dental education. *Finally,* because NM assumes an elimination-and-guessing strategy, it may not fully capture complex student reasoning processes such as partial knowledge or

misconceptions. A further limitation is the absence of item-level psychometric data. Consequently, the passing score derived from expert judgement could not be evaluated alongside observed indicators of item difficulty or discrimination.

6. CONCLUSION

The application of NM in this study provided a structured approach to determining a passing score for the Oral and Maxillofacial Pathology 1 examination. Compared with fixed passing scores, this method established a passing standard based on item-level analysis and expert judgment, helping ensure that the passing threshold better reflects the expected competence of dental students. The results also showed that criterion-referenced standard setting can help identify borderline students who might otherwise pass under traditional rules. This is particularly important in health professions education, where assessment decisions are closely related to clinical competence and patient safety.

Overall, the findings support the use of evidence-based standard-setting methods such as the Nedelsky approach to improve the transparency, fairness, and validity of assessment practices in dental education.

DECLARATIONS

Funding: This research is funded by Van Lang University, Vietnam under grant number VLU-2503-DT-KRHM-GV-0072.

Institutional Review Board Statement: The research protocol was reviewed and approved by the Research ethics Committee of Van Lang University (No. 27/2026/IRB-VN01.078; Date of approval: March 18th, 2026).

Transparency: The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted.

Competing Interests: The authors declare that they have no competing interests.

Authors' Contributions: H.M.V.: research design, methodology, writing original draft, manuscript preparation; T.T.H.N.: methodology, writing original draft, manuscript preparation; N.T.P.P.: writing original draft, manuscript preparation; L.T.N.: writing original draft, manuscript preparation; V.H.A.D.: research design, methodology, writing original draft, data analysis, manuscript preparation. All authors have read and agreed to the published version of the manuscript.

Disclosure of AI Use: Artificial intelligence tools were utilized exclusively for minor language polishing (grammar checking and vocabulary refinement) and not for generating scientific ideas, content, or analyzing data.

REFERENCES

Abd-Rahman, A. N. A., Baharuddin, I. H., Abu-Hassan, M. I., & Davies, S. J. (2021). A comparison of different standard-setting methods for professional qualifying dental examination. *Journal of Dental Education, 85*(7), 1210-1216. https://doi.org/10.1002/jdd.12600

Albino, J. E., Young, S. K., Neumann, L. M., Kramer, G. A., Andrieu, S. C., Henson, L., Horn, B., & Hendricson, W. D. (2008). Assessing dental students' competence: Best practice recommendations in the performance assessment literature and investigation of current practices in predoctoral dental education. *Journal of Dental Education, 72*(12), 1405-1435. https://doi.org/10.1002/j.0022-0337.2008.72.12.tb04632.x

Bahammam, L. A. (2017). Cutoff score evaluation of undergraduate dental curriculum. *Medical Teacher, 39*(sup1), S33-S36. https://doi.org/10.1080/0142159X.2016.1254741

Chang, L. (1999). Judgmental item analysis of the Nedelsky and Angoff standard-setting methods. *Applied Measurement in Education, 12*(2), 151-165. https://doi.org/10.1207/s15324818ame1202_3

Cohen-Schotanus, J., & van der Vleuten, C. P. (2010). A standard setting method with the best performing students as point of reference: Practical and affordable. *Medical Teacher, 32*(2), 154-160. https://doi.org/10.3109/01421590903196979

Dalum, J., Christidis, N., Myrberg, I. H., Karlgren, K., Leanderson, C., & Englund, G. S. (2023). Are we passing the acceptable? Standard setting of theoretical proficiency tests for foreign-trained dentists. European journal of dental education : official journal of the *Association for Dental Education in Europe, 27*(3), 640-649. https://doi.org/10.1111/eje.12851

Downing, S. M., Tekian, A., & Yudkowsky, R. (2006). Research methodology: Procedures for establishing defensible absolute passing scores on performance examinations in health professions education. *Teaching and Learning in Medicine, 18*(1), 50-57. https://doi.org/10.1207/s15328015tlm1801_11

Downing, S.M., & Yudkowsky, R. (2009). *Assessment in health professions education* (1st ed.). Routledge. https://doi.org/10.4324/9780203880135

George, S., Haque, M. S., & Oyebode, F. (2006). Standard setting: comparison of two methods. *BMC Medical Education, 6*, 46. https://doi.org/10.1186/1472-6920-6-46

Ho, T. K., Abu Kassim, N. L., O'Malley, L., & Roudsari, R. V. (2025a). Standard setting for dental knowledge tests: Reproducibility of the modified Angoff and Ebel method across judges. *BMC Medical Education, 25*, 1426. https://doi.org/10.1186/s12909-025-07822-3

Ho, T. K., O'Malley, L., & Roudsari, R. V. (2025b). Investigating assessment standards and fixed passing marks in dental undergraduate finals: A mixed-methods approach. BMC *Medical Education, 25*(1), 481. https://doi.org/10.1186/s12909-025-06944-y

Kim, J., & Yang, J. S. (2020). How to improve reliability of cut-off scores in dental competency exam: A comparison of rating methods in standard setting. *European Journal of Dental Education, 24*(4), 734-740. https://doi.org/10.1111/eje.12563

Mortaz Hejri, S., & Jalili, M. (2014). Standard setting in medical education: fundamental concepts and emerging challenges. *Medical journal of the Islamic Republic of Iran, 28*, 34.

Nedelsky, L. (1954). Absolute grading standards for objective tests. *Educational and Psychological Measurement, 14*(1), 3-19. https://doi.org/10.1177/001316445401400101

Norcini, J. J. (2003). Setting standards on educational tests. *Medical Education, 37*, 464-469. https://doi.org/10.1046/j.1365-2923.2003.01495.x

Shulruf, B., Wilkinson, T., Weller, J. et al. Insights into the Angoff method: results from a simulation study. *BMC Med Educ 16*, 134 (2016). https://doi.org/10.1186/s12909-016-0656-7

Verhoeven, B. H., Verwijnen, G. M., Muijtjens, A. M., Scherpbier, A. J., & van der Vleuten, C. P. (2002). Panel expertise for an Angoff standard setting procedure in progress testing: item writers compared to recently graduated students. *Medical education, 36*(9), 860-867. https://doi.org/10.1046/j.1365-2923.2002.01301.x

Yousef, M. K., Alshawwa, L., Tekian, A., & Park, Y. S. (2017). Challenging the arbitrary cutoff score of 60%: Standard setting evidence from preclinical Operative Dentistry course. *Medical teacher, 39*(sup1), S75-S79. https://doi.org/10.1080/0142159X.2016.1254752

Yousef, M. K., Alshawwa, L. A., Farsi, J. M., Tekian, A., & Yoon Soo, P. (2020). Determining defensible cut-off scores for dental courses. *European journal of dental education : official journal of the Association for Dental Education in Europe, 24*(2), 186-192. https://doi.org/10.1111/eje.12483