Alqalam Journal of Medical and Applied Sciences. 2025;8(4):2720-2727
https://doi.org/10.54361/ajmas.258489

Original article

A Structured Approach to Assessing Clinical Competence in Conservative Dentistry: A Checklist-Based Evaluation for Dental Students performing Class I Composite Restorations

Khaled Elhenshiri*, Mohamed Benhakoma, Fatma Soltan, Asma Abognah, Farida Alsayeh

Department of Endodontic and Conservative Dentistry, Faculty of Dentistry, University of Tripoli, Tripoli, Libya.
Corresponding Email. kh.elhenshiri@uot.edu.ly

Abstract
Accurate evaluation of clinical competence in restorative dentistry requires reliable assessment tools. Checklists - based standardized tool has the potential to enhance objectivity, but not full representation of students' underlying knowledge. This study aimed to validate a structured checklist for Class I composite restorations by examining inter-examiner reliability and comparing it with novice students' self-assessment accuracy. The cross-sectional study involved thirty undergraduate dental students. Each student performed a standardized Class I composite restoration, assessed concurrently by three independent faculty examiners and through student self-evaluation using the same instrument. Internal consistency, Cronbach's (α), inter-examiner reliability (ICC), and Pearson correlation (r) were used for analysis. The checklist showed good internal consistency (α = 0.83) and moderate agreement amongst examiners (ICC = 0.60). The mean score for student self-assessment (33.6) was slightly higher than the mean score for examiner assessments (32.7), but their relationship was negligible and negative (r = -0.06, p = 0.75). Significantly, including the self-assessment scores reduced the overall reliability of the scale, with the ICC dropping sharply to 0.30, indicating low correspondence. The developed checklist proves to be a reliable and objective measure for clinical assessment. These findings, however, emphasize a critical need for enhanced examiner calibration and integration of formal reflective practice training to improve students' self-assessment accuracy and professional self-awareness.
Keywords. Clinical Competence, Checklist, Self-Assessment, Reliability, Dental Education.

Introduction

Clinical competence is one of the bases of competency-based dental education, and the goal is to graduate student practitioners who can practice independently. It encompasses the integration of theoretical knowledge, practical skills, and professional attitudes that collectively define competent performance in clinical settings [1,2]. Because theoretical knowledge alone does not ensure sound clinical judgement, the assessment of competence must also capture reasoning and decision-making in real-world contexts [3]. Within dental education, structured assessments such as the Objective Structured Clinical Examination (OSCE) are widely used to evaluate students' attainment and development of these clinical competencies [4,5]. This proficiency, built on a holistic structure of knowledge and technical skills, includes core components like Class I composite restorations, which embody skills typically required in contemporary practice [6,7]. To ensure students possess the desired standard of proficiency, reliable and verifiable assessment frameworks should be present [3,7].

Standard clinical evaluation methods, often dependent on global ratings or personal opinion, have been criticized as being unreliable and prone to testing bias [5,7]. The consequence is non-standardized procedures more often than not yielding unequal marking and less credibility, with a lack of consistent feedback and persistent inter-examiner variation [6,8]. This dilemma is also compounded by the inherent technical difficulties of contemporary procedures like composite restorations, which require precise manipulation to manage factors like polymerization shrinkage, challenging marginal adaptation, and moisture control [6,9]. In addition, the mandatory application of student self-assessment—a practice that is supposed to facilitate reflective learning—has its own pitfall: it is susceptible to inconsistency and bias stemming from student inexperience, anxiety, or tension when performing such complex clinical procedures for the first time on patients [10–12]. This inherent variability, in conjunction with persistent disagreements among external examiners, necessitates more formalized and open assessment protocols with the potential to provide valid and reliable measurement of complex clinical skills [5,7,13].

Hence, checklist-type assessment tools have come as a successful solution to such limitations. With segmentation of procedures into systematically defined steps with criteria that can be defined, checklists enable criterion-referenced assessment, thereby enhancing the objectivity of scores but at the same time decreasing inter-examiner variation and enhancing consistency of clinical evaluation [8,13,14,16]. The tools also provide students with tangible, concrete results upon which they can maximize clinical competence building [12,13,17].

Therefore, the present study employs a systematic checklist to not only measure inter-examiner reliability but also to specifically examine the discrepancy between expert examiner scores and potentially biased student self-ratings for Class I composite restorations.

Received*: 24-09-2025 -* ***Accepted****: 22-11-2025 -* ***Published****: 30-11-2025*

Methodology

Study Design

The cross-sectional observational study assessed the reliability and internal consistency of the Class I cavity preparation and composite restoration checklist filled by undergraduate dental students at the Faculty of Dentistry, University of Tripoli. The key objectives were to test inter-rater reliability and the comparison of the examiners' assessment against students' self-evaluations under examination conditions.

Participants

Fifty-four (54) fourth-year dental students enrolled in the Conservative Dentistry clinical department during the 2024–2025 academic year at the Faculty of Dentistry, University of Tripoli. The inclusion criterion required that each student, as part of the standard clinical curriculum, had already performed a specified number of Class I cavity preparations and composite restorations on patients.

Twenty-four (24) students who did not meet these fundamental requirements were removed, leaving a final study sample of thirty (30) students. All thirty (30) students performed one Class I cavity preparation and composite restoration on a natural tooth as part of their routine clinical examination. Three examiners who undertook the assessment had more than eight years of specialized clinical teaching and practice experience. Each of them independently assessed the 30 student performances. All students who participated in this research signed a consent form to participate. Ethical approval was also obtained from the Libyan National Committee for Biosafety and Bioethics (Ref N®: NBC: 002.H-25.34). Confidentiality and anonymity were maintained at all steps throughout the data collection process.

Checklist Development and Validation

The criterion-referenced, standardized checklist was developed by merging two existing operative dentistry checklists and was subsequently refined and modified using expert review and content analysis for validation. Item-level Content Validity Index (I-CVI) was calculated from expert ratings on representativeness, relevance, and clarity. The Scale-level Content Validity Index (S-CVI) was determined as the mean across all items. Items with I-CVI < 0.70 were modified or replaced based on the model proposed by Lynn M.R. (1986) [18]. The ultimate checklist was reduced to 20 items, divided into four domains:(1) infection control, chair position, and anesthesia administration; (2) Rubber dam placement and cavity preparation; (3) pre-restorative protocol (etching and bonding technique); (4) restorative protocol. Student performance was scored for each item on a three-point scale (0 = Unsatisfactory, 1 = Satisfactory, 2 = Excellent), with a maximum possible total score of 40. The final checklist is given in (Table 1).

Table 1. Final checklist evaluation for Class I Composite Restorations

Operative Dentistry Checklist				
I. Infection Control, Chair Position, and Anesthesia Administration				
Domain /Criteria	**2 = excellent**	**1 = satisfactory**	**0 = unsatisfactory**	**score**
Student dressed in a proper clean uniform and PPE	☐	☐	☐	
All instruments were sterilized and arranged aseptically	☐	☐	☐	
Disinfection of the working area and equipment was performed	☐	☐	☐	
Were the patient and student positioned properly?	☐	☐	☐	
Pain-free administration of local anesthesia and effectiveness before starting	☐	☐	☐	
II. Rubber Dam Application and Cavity Preparation:				
Appropriate clamp selection without trauma to the gingiva	☐	☐	☐	
Stable and moisture-proof isolation achieved	☐	☐	☐	
Conservative cavity outlines respecting caries extent	☐	☐	☐	
Complete caries removal without overcutting	☐	☐	☐	
Dry, contamination-free cavity before restoration	☐	☐	☐	
III. Pre-Restorative protocol (Etching and bonding technique)				
Acid Etchant applied properly: duration, and coverage (e.g., selective, total)	☐	☐	☐	
Etched surface rinsed and dried properly without over-drying dentin	☐	☐	☐	

Received*: 24-09-2025 -* ***Accepted****: 22-11-2025 -* ***Published****: 30-11-2025*

Bonding agent applied uniformly with appropriate technique (e.g., scrubbing)	☐	☐	☐	
Solvent evaporated completely before light curing	☐	☐	☐	
Light curing was performed for the correct duration and distance	☐	☐	☐	
IV. Restorative Protocol:				
Pulp protection, when necessary, using both traditional and biomimetic methods	☐	☐	☐	
Composite is applied in appropriate increments, with curing performed for each increment	☐	☐	☐	
Assessing premature contacts and adjustments (Articulating paper)	☐	☐	☐	
Final contouring and occlusal anatomy were accurately restored	☐	☐	☐	
Margins well-finished and Sealed	☐	☐	☐	

Assessment Protocols and Examiner Calibration

In controlled clinical conditions, all 30 students prepared and completed a Class I cavity preparation and composite restoration in a natural tooth. The assessment was done under a strict protocol for independent assessment. In the clinical setup, the first evaluator directly observed and assessed the performance of every student based on the final checklist and recorded scores as (ES1). For the purpose of determining inter-examiner reliability, two other examiners utilized the identical checklist to independently and blindly score the same procedures, and their scores were recorded as (ES2) and (ES3) for the identical students, respectively. In order to ensure the validity and reliability of data gathered, standardization was employed to ensure reliability among examiners for each step of data gathering. In addition to the examiner scores, immediately after completion of the clinical procedure, a final checklist was distributed to each student, who was then asked to complete it as a self-evaluation of their own performance. The students' self-evaluation scores (SS) were recorded for comparison with expert examiner scores. This comprehensive approach allowed not only the measurement of inter-examiner reliability but also an assessment of the degree to which students' self-evaluations of their performance approximated expert opinion.

Data Analysis

Four sets of scores were recorded for each clinical case: three examiner scores (ES1, ES2, ES3) and one student score (SS). All data were anonymized using coded identifiers, which were then kept in a password-protected database. Both item-level scores (0–2) and total scores (0–40) for all raters were included in the dataset. All statistical analyses were conducted using DATAtab online software (DATAtab e.u., Graz, Austria). Descriptive statistics were first computed to summarize the distribution of student performance scores, examiner ratings, and self-assessment scores. Measures of central tendency (mean, median) and variability (standard deviation, range) were reported to provide an overview of the data. The assumption of normality was evaluated prior to inferential testing using both graphical methods (i.e., histograms) and statistical tests of normality (i.e., Kolmogorov–Smirnov test) to ensure the appropriateness of subsequent parametric analyses.

To assess the internal consistency of the checklist, Cronbach's alpha (α) was calculated across all items. A coefficient value of ≥ 0.70 was considered indicative of acceptable internal reliability. Reliability analysis determined the level of agreement among the three examiners: Inter-examiner reliability was assessed using the Intra-class Correlation Coefficient (ICC), applying a Two-way Random Effects Model with absolute agreement. Comparisons between examiner scores and student self-scores were analyzed using the Pearson correlation coefficient (r) was calculated to quantify the linear relationship between student self-assessment scores (SS) and average examiner score (Avg ES). Bland–Altman plots were employed to evaluate the agreement between student self-assessment scores and the average scores of the three examiners. This method is particularly useful for determining whether two measurement methods can be used equivalently. Statistical significance was set at $p < 0.05$.

Results

Descriptive Statistics

Across the three examiners in (Table 2), mean scores were closely aligned, ranging from 32.1 to 33.39, with standard deviations between 1.86 and 2.98. Examiner 2 tended to award slightly higher and less variable scores, while Examiners 1 and 3 showed wider ranges. The overall mean score was 32.7 (with values spanning 22 to 36; SD = 2.44; 95% CI: 31.79–33.61). These findings indicate consistent examiner evaluations, with most students achieving satisfactory to high performance levels on the checklist.

Received*: 24-09-2025 -* ***Accepted****: 22-11-2025 -* ***Published****: 30-11-2025*

Table 2: Descriptive Statistics for Examiner and Student Assessment Scores

Descriptive Statistics	Examiner 1	Examiner 2	Examiner 3	Self-assess student	Average scores of three examiners
Mean	32.1	33.39	32.55	33.6	32.7
Std. Deviation	2.98	1.86	2.9	2.24	2.44
Minimum	22	29	25	27	26
Maximum	36	36	36	36	36
95% Confidence interval for mean	31 - 33.19	32.71 - 34.07	31.49 - 33.61	32.76 - 34.44	31.79 - 33.61

The Kolmogorov–Smirnov test indicated that the examiner and student scores did not deviate significantly from normality (Figure 1), with p-values greater than 0.05.

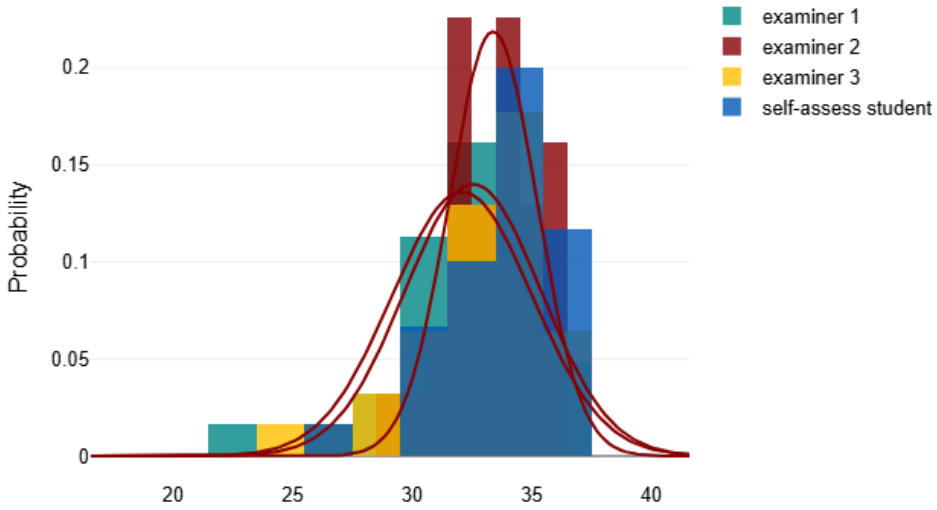

Figure 1: Probability Distribution of Assessment Scores Across Examiners and Students

The assessment scores generated by three professional examiners (ES1, ES2, and ES3) and the self-assessment scores provided by the students (SS) were demonstrated using a box-and-whisker plot (Figure 2).

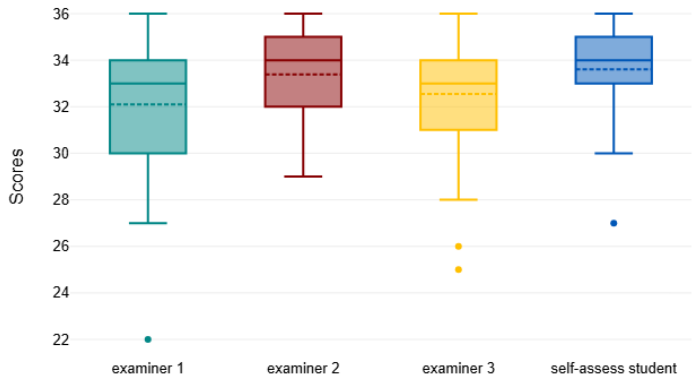

Figure 2. Distribution of Assessment Scores by Professional Examiners and Self-Assessing Students

A marked difference in the central tendency of the scores was observed across the four groups. The Self-Assess Student group and Examiner 2 recorded the highest median scores, both approximately 34. In contrast, Examiner 1 and Examiner 3 had lower median scores of approximately 33 and 32, respectively. The mean scores (indicated by the dashed line) followed a similar pattern, with the Self-Assess Student group demonstrating the highest overall average score (approximately 33.8).

The interquartile range (IQR) indicated the spread of scores for each group (represented by the box length). Examiner 2 exhibited the narrowest IQR (3 points, from 32 to 35), suggesting the highest level of agreement among scores within the central 50% of that group. Examiner 1 and Examiner 3 displayed a greater score spread within the IQR, particularly Examiner 3, whose box stretched from 31 to 34. The Self-Assess Student

Received*: 24-09-2025 -* ***Accepted****: 22-11-2025 -* ***Published****: 30-11-2025*

group also showed a relatively narrow IQR (from 33 to 35), indicating a tight clustering of the majority of self-reported scores, consistent with their high median score.
The presence of outliers suggests scores that deviate significantly from the rest of the distribution. Examiner 1 and the self-assess student group each presented one distinct outlier on the lower end, at scores of approximately 22 and 27, respectively. Examiner 3 showed the highest number of low-end outliers, with two scores recorded near 26 and 25. Examiner 2 was the only examiner to show no visible outliers, with all scores falling within the whisker range, further indicating the consistency of their scoring. The self-assessed student scores clustered tightly and skewed toward the upper range, comparable to the most generous professional assessor (Examiner 2), while Examiner 3 reported the lowest central tendency and the highest number of low-end outliers.

Internal Consistency and Reliability

The internal consistency of examiner scores was assessed using Cronbach's alpha (α). The analysis yielded an α of 0.83 across the three examiners, indicating good reliability. Corrected item–total correlations were strong for Examiner 1 (0.65), Examiner 2 (0.69), and Examiner 3 (0.84). Deleting any examiner's ratings did not improve reliability, confirming that all examiners contributed positively to the overall consistency of the scale. When student self-assessment scores were included, Cronbach's alpha (α) decreased to 0.65 (indicating only questionable internal consistency), and the corrected item–total correlation for the self-assessment score was negative (-0.05), suggesting poor alignment with examiner ratings. Removing the student self-scores immediately increased Cronbach's alpha (α) to 0.84, confirming that student self-scores reduced the overall internal consistency of the scale.

Inter-Examiner Agreement and Correlation

Inter-examiner reliability was evaluated using the Intraclass Correlation Coefficient (ICC). The analysis yielded an ICC of 0.60 (95% CI: 0.39 - 0.76), indicating moderate agreement among the three examiners. The associated F-test was statistically significant ($F = 6.0$, $p < 0.001$), suggesting the reliability was not due to chance. This finding demonstrates that examiners achieved a consistent, though moderate, level of agreement. When student self-assessment scores were included, the ICC (for the 4-rater group) decreased significantly to 0.30 (95% CI: 0.13 – 0.51), confirming limited consistency with examiner evaluations. Furthermore, Pearson's correlation showed a negligible negative relationship between student self-assessment scores (SS) and the average examiner scores (Avg ES) ($r(28) = -0.06$, $p = 0.75$). This result indicates that student self-evaluations did not correspond with the examiner's assessments.
The Bland–Altman plot confirmed that the mean difference between self-assessment scores and examiner scores was 0.9, indicating that students, on average, scored themselves slightly higher than the examiners (Figure 3). The 95% limits of agreement ranged from -5.78 to 7.58, demonstrating wide variability between the two scoring methods. Several data points fall close to or beyond these limits, suggesting inconsistent alignment between student self-evaluations and examiner assessments. Overall, the analysis indicates that self-assessment scores cannot be considered interchangeable with examiner ratings.

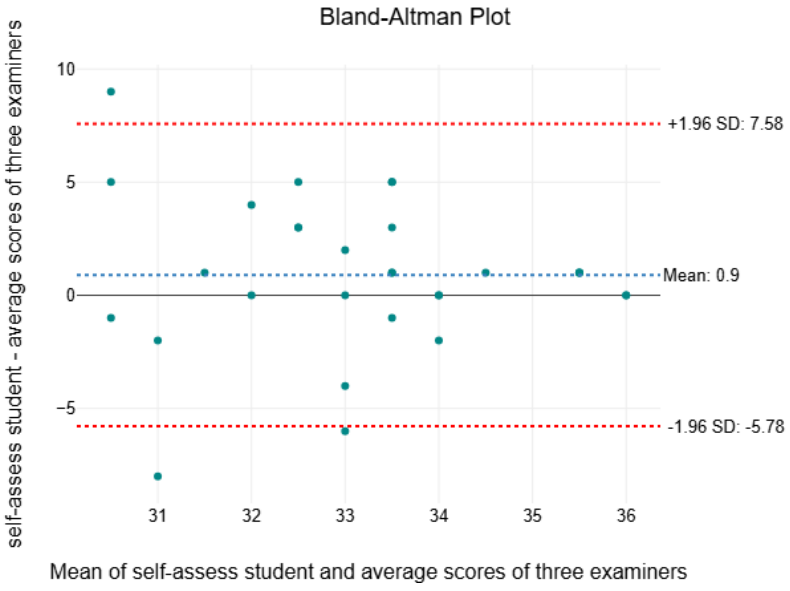

Figure 3. The Agreement Between Student Self-Assessment and Average Examiner Scores

These findings are consistent with the ICC and Pearson correlation results, all of which demonstrated low agreement between student self-assessments and examiner evaluations.

Received*: 24-09-2025 -* ***Accepted****: 22-11-2025 -* ***Published****: 30-11-2025*

Discussion

The study gauged the reliability of the examiners and the students when using a checklist structured for Class I composite restorations in a clinical practice. The narrow range of means for the examiners (32.1 – 33.39) and the overall mean score (32.7 ±2.44) indicate close agreement among assessments. A Cronbach's (α = 0.83) indicates strong internal consistency, confirming that checklist-based evaluations can be used as dependable measures of clinical competence. These findings align with those of O'Donnell et al., who emphasized the value of rubrics in promoting defensible and transparent grading, and Daghrery et al., who reported moderate to high inter-rater reliability in Class II restoration assessments using analytic rubrics [16,19]. The inter-examiner reliability of the current study (ICC = 0.60) is in agreement with prior literature showing that systematic assessment tools and calibration reduce—but do not eliminate—clinician subjectivity in grading. While α = 0.83 attests to the tool's internal quality, the moderate ICC = 0.60 advocates that the inherent limitations of structured tools persist. This challenge was noted by Vann and Machen, who argued that checklists may not always surpass global methods in achieving optimal inter-rater agreement [17]. Slight improvement in internal consistency after excluding student self-assessment (α = 0.84) suggests that examiner ratings are a more reliable measure of performance. This finding aligns with Mittal et al., who observed moderate correlations (r ≈ 0.7) between student and staff ratings regardless of the students' overall competence level [12].

The negative and non-significant correlation (r = –0.06, p = 0.75) between examiner and student ratings supports the conclusion that students tend to overestimate their own performance. The Bland–Altman plot supported this, showing a mean difference of (0.9) points (students rating themselves higher) and wide limits of agreement (–5.78 to 7.58). This finding is consistent with earlier research by Satterthwaite and Grey, who reported only moderate agreement (Kappa = 0.32–0.53) between peer ratings and expert assessment in operative dentistry [14]. These findings reflect a well-known cognitive bias in self-assessment, where limited clinical experience often leads to overconfidence and, consequently, reduced accuracy in self-evaluation. The moderate inter-examiner reliability (ICC) found in this study aligns with earlier research that highlights variation among raters. Park et al. demonstrated that faculty rank and individual judgment can influence assessment outcomes, despite the application of standardized criteria [15]. This underscores the need for examiner calibration, professional development, and routine review of assessment criteria to maximize inter-examiner reliability. Furthermore, as Bilan and Negahdari observed, institutional barriers—such as inconsistent faculty engagement and limited resources—often hinder the effective implementation of performance assessments in dental education [20].

These results confirmed that list-based and organized assessments are valuable components of competency-based learning environments. Structured checklists and analytical rubrics ensure defensibility, objectivity, and transparency in grading, which are essential factors in clinical education [17,20,21]. Moreover, formative feedback and repeated assessments are more effective than single summative assessments in fostering learning and self-awareness [12,13,22]. However, the prescriptive accuracy of checklists can, unintentionally, serve as a blind catalyst for "teaching to the checklist," whereby students are more worried about accomplishing rubric elements than developing more robust clinical thinking. Holmboe et al. cautioned that such a form of assessment reductionism could, in turn, overemphasize procedural conformity at the expense of real clinical capability [13]. In a similar vein, Durning et al. emphasized that authentic assessment should capture reasoning *in vivo*—within real clinical contexts—rather than *in vitro* through overly standardized measures [3]. Similarly, van der Vleuten and Lambert argued that rigid standardization can potentially limit ecological validity, thus potentially deterring adaptive response in patient-specific contexts [23]. These concerns are echoed by Albino et al., who advocate for performance assessments that balance structure with flexibility to reflect real-world competence [24].

This reduction in trust in self-assessment highlights a significant pedagogical challenge: building dependable self-evaluation skills in clinical education. Khanghahi and Azar's reviews indicate that accuracy in self-assessment rises with guided feedback and structured reflective practice [22]. Therefore, student self-assessment must supplement—not replace—faculty assessment in competency-based education.

Criticisms in the evaluation literature frequently identify the limitations of decomposing complex clinical competencies into discrete checklist items. Cook et al. remind us that validity needs to be viewed as an accumulation of evidence over content, internal structure, relationships to other variables, and consequences [25]. Over-reliance on discrete checklists risks oversimplifying complex constructs and excluding the tacit knowledge and adaptive decision-making essential to real-world clinical performance.

These observations align with the broader principles of competency-based education, which advocate for flexible, criterion-referenced systems, as described by Albino et al. and ElBadrawy and Korayem [24, 26]. Standardized checklists, when paired with repeated examiner calibration, offer a stable and highly structured platform for assessing clinical skills. Nevertheless, complete objectivity remains elusive, and excessive reliance on rigid checklists may unintentionally narrow the educational focus—potentially limiting the development of clinical reasoning and adaptive decision-making.

Received*: 24-09-2025 -* ***Accepted****: 22-11-2025 -* ***Published****: 30-11-2025*

Conclusion

This study demonstrated that adopting a structured, checklist-based approach provides a reliable and systematic method for evaluating Class I composite restorations in dental education. The strong internal consistency (α = 0.83) observed among examiners confirms that the checklist functions as a robust, criterion-referenced assessment instrument. However, two persistent challenges were identified: First, the moderate inter-examiner reliability (ICC = 0.60), and second, the negligible correlation (r = –0.06) between examiner and student scores. These findings underscore the need for continuous examiner calibration and structured self-assessment training. Integrating reflective feedback within clinical instruction is therefore essential to foster both accuracy and professional growth. Ultimately, reliable assessment tools must not only ensure grading fairness but also foster the development of self-awareness and clinical judgment—the cornerstones of professional competence in dentistry.

Acknowledgments

The authors wish to express their gratitude to all participants for their valuable contributions, which made this work possible.

Conflicts of interest

None.

References

1. Albino JEN, Young SK, Neumann LM, Kramer GA, Andrieu SC, Henson L, et al. Assessing dental students' competence: best practice recommendations in the performance assessment literature and investigation of current practices in predoctoral dental education. J Dent Educ. 2008 Dec;72(12):1405-35.
2. Bilan N, Negahdari R. The competency-based evaluation of educational crew of dental faculty's obstacles in institutionalizing performance assessments. Open Dent J. 2022 Jan 18;16(1):e187421062201040.
3. Christensen GJ. Remaining challenges with class II resin-based composite restorations. J Am Dent Assoc. 2007 Nov;138(11):1487-9.
4. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. Med Educ. 2015 Jun;49(6):560-75.
5. Daghrery A, Alwadai GS, Alhussein AS, Alzahrani MM, Almana MA, Bhavikatti SK, et al. Students' performance in clinical class II composite restorations: a case study using analytic rubrics. BMC Med Educ. 2024 Nov 14;24(1):1252.
6. de Souza FF, Barros I, Chisini LA, Cademartori MG, Costa FDS, Demarco FF. Stress amongst dental students in the transition from preclinical training to clinical training: a qualitative study. Eur J Dent Educ. 2023 Aug;27(3):568-74.
7. Durning SJ, Holmboe ES. Assessing clinical reasoning: Moving from in vitro to in vivo. Diagnosis (Berl). 2014;1(1):111-7.
8. ElBadrawy HE, Korayem M. The flexible requirement system for grading of clinical performance of undergraduate dental students. Eur J Dent Educ. 2007 Nov;11(4):208-15.
9. Hallak JC, Ferreira FS, Mello BZ, Demarco FF, Oling J, Oenning ACC. Transition between preclinical and clinical training: perception of dental students regarding the adoption of ergonomic principles. PLoS One. 2023 Mar 16;18(3):e0283084.
10. Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. The role of assessment in competency-based medical education. Med Teach. 2010;32(8):676-82.
11. Khanghahi ME, Azar FE. Direct observation of procedural skills (DOPS) evaluation method: systematic review of evidence. Med J Islam Repub Iran. 2018 May 29;32:45.
12. La Chimea T, Kanji Z, Rambihar V. Assessment of clinical competence in competency-based education. Can J Dent Hyg. 2020 Jun 1;54(2):83-91.
13. Lynn MR. Determination and quantification of content validity. Nurs Res. 1986 Nov-Dec;35(6):382-5.
14. Manogue M, Brown G. Clinical assessment of dental students: values and practices of teachers in restorative dentistry. Med Educ. 2001 Apr;35(4):364-70.
15. Mittal P, Jadhav GR, Singh A, Singh S, Kumar P, Kaur H. Impact of self–assessment on dental student's performance in pre-clinical conservative dentistry course. BMC Oral Health. 2024 May 29;24(1):593.
16. Moura FR, Romano AR, Lund RG, Piva E, Rodrigues Júnior SA, Demarco FF. Three-year clinical performance of composite restorations placed by undergraduate dental students. BMC Oral Health. 2011 Feb 7;11:3.
17. Naseer A, Aziz S, Khan AM, Nazeer M, Afzal A, Ahmad S. Analysis of satisfaction levels and perceptions of clinical competency: a mixed method study on objective structured clinical examination in undergraduate dental students. BMC Med Educ. 2024 Jun 18;24(1):673.
18. O'Donnell JA, Oakley M, Haney S, O'Neill PN, Taylor D. Rubrics 101: a primer for rubric development in dental education. J Dent Educ. 2011 Sep;75(9):1163-75.
19. Park RD, Susarla SM, Cox CK, Da Silva J, Howell TH. Differences in clinical grading associated with instructor status. Eur J Dent Educ. 2009 Feb;13(1):31-8.
20. Rayyan MR. The use of objective structured clinical examination in dental education – a narrative review. Front Oral Health. 2024 May 9;5:1336677.
21. Satterthwaite JD, Grey NJ. Peer-group assessment of pre-clinical operative skills in restorative dentistry and comparison with experienced assessors. Eur J Dent Educ. 2008 May;12(2):99-102.
22. Smales RJ, Yip HK. Competency-based education in a clinical course in conservative dentistry. Br Dent J. 2001 Nov 10;191(9):517-22.

Received*: 24-09-2025 -* ***Accepted****: 22-11-2025 -* ***Published****: 30-11-2025*

23. Taleghani M, Solomon ES, Wathen WF. Non-graded clinical evaluation of dental students in a competency-based education program. J Dent Educ. 2004 Jun;68(6):644-55.
24. van der Vleuten CP, Schuwirth LW, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, et al. A model for programmatic assessment fit for purpose. Med Teach. 2012;34(3):205-14.
25. van der Vleuten CP, Schuwirth LW. Assessment in the context of problem-based learning. Adv Health Sci Educ Theory Pract. 2019 May;24(5):903-914.
26. Vann WF Jr, Machen JB. Effects of criteria and checklists on reliability in preclinical evaluation. J Dent Educ. 1983 Oct;47(10):671-5.

Received*: 24-09-2025 -* ***Accepted****: 22-11-2025 -* ***Published****: 30-11-2025*