The Saudi Dental Journal 36 (2024) 1253–1259

Contents lists available at ScienceDirect

The Saudi Dental Journal

journal homepage: www.ksu.edu.sa
www.sciencedirect.com

Original Article

Comparing the reliability of inter- and intra-grader using digital scanning vs. traditional visual method for evaluating preclinical class II composite preparation

Amr Ahmed Azhari [a], Walaa Magdy Ahmed [a,*], Raghad Almaghrabi [b], Aroob Almalki [b], Yasser Merdad [a], Hind J. Alrefai [c], Mawadah T. Bahrawi [c], Lama Faisal Kordi [c]

[a] Department of Restorative Dentistry, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia
[b] Internship Program, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia
[c] Department of Conservative Dentistry, University Dental Hospital, King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Keywords:
Visual grading methods
Digital grading methods
Intraoral scanners
Class II preparations
Intra-grader reliability

ABSTRACT

Objective: This study aimed to evaluate the intra- and inter-grader reliability of four evaluators using three different digital intraoral scanners and visual methods for typodontic Class II composite preparations.
Materials and methods: Ninety-five typodont teeth of Class II composite preparations were evaluated using traditional visual grading methods (VGM) and digital grading methods (DGM) using the same rubric. Three intraoral scanners were used to scan the Class II cavity preparation for the composite: i700 (Medit, Korea), Trios 4 (3Shape, Denmark), and Shinning 3D (Shinning 3D, China). The same rubric was used to score the visual and digital evaluations by calibrated examiners. Two-way ANOVA was used to compare method- and evaluator-based scores, accounting for the scanner type used.
Results: The scores of the prepped typodont teeth were subjected to an interaction between the examiner and the evaluation technique. In addition, the mean total prepped teeth scores differed between examiners using VGM. A statistically significant interaction emerged between examiners and the evaluation technique employed to assess the total score of the prepped teeth: $F(9, 1504) = 3.893$, $P = 0.001$, partial $\eta 2 = 0.023$. The total prepped tooth score differed between the VGM and DGM groups. Lower ($P < 0.05$) intra-grader consistency was observed for the final scores when Class II preparations were evaluated using the VGM; however, this consistency improved when using the DGM.
Conclusion: Examiners and evaluation methods affect student performance in Class II cavity preparations. The DGM may be more reliable and consistent within and between evaluators than the VGM is.

1. Introduction

Predictability, efficiency, and accuracy have increased with the increasing use of digital dentistry tools. Concurrently, demand for the integration of digital technology into dental education has increased (Stoilov et al., 2021). Technology-based educational systems involve computer and computer-assisted simulations. Currently, there are two main types of digital dental education systems. One is a force-feedback-based simulator that uses a haptic device and virtual human teeth in oral cavities as a platform to enable the practice of dental procedures in a simulated environment. The other one is a digital evaluation system that provides feedback on the student's preparation.

Preclinical assessment are essential components in the development of dental students. However, traditional methods, such as visuals, lack consistency and objectivity. Digital tools provide precise feedback, support student development, and improve performance before clinical practice (Renne et al., 2013). The fine motor skills and hand-eye coordination required to perform specific dental procedures may improve when the procedure is repeated in a safe environment and graded objectively using defined criteria for evaluation (Clancy et al., 2002). According to Stoilov et al. (2021), digital assessments of preclinical dental work appear to be comparable to conventional visual feedback. However, they can help improve students' learning curves, promote independent learning, reduce faculty overload, and improve students'

* Corresponding author at: Department of Restorative Dentistry, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia.
E-mail address: wmahmed@kau.edu.sa (W.M. Ahmed).

https://doi.org/10.1016/j.sdentj.2024.07.005
Received 23 December 2023; Received in revised form 30 May 2024; Accepted 4 July 2024
Available online 6 July 2024

self-assessment of cavity preparation (Kim et al., 2023).

A study explored the use of digital assessments for educational purposes and highlighted their potential in preclinical and clinical settings (Jorquera et al., 2021). Several studies have compared the effectiveness of digital assessment tools in evaluating clinical procedures using traditional visual assessment methods (Tiu et al., 2016; Nagy et al., 2018; Greany et al., 2019; Miyazono et al., 2019). Studies have shown that digital assessments are more accurate than visual inspections (Lam et al., 2015; Miyazono et al., 2019). Others have reported that digital evaluation technology has no impact compared to visual inspection (Gratton et al., 2016; Greany et al., 2019) or stated that digital assessment could not be as effective for student learning as narrative feedback (Gratton et al., 2017). Lam et al. (2015) showed that digital assessment correlated well with visual scores, offering consistent results even when multiple scanners were used.

One study compared conventional evaluation methods with digitally scanned assessments of Class II amalgam and III resin composite student cavity preparations. Both methods showed no difference in evaluation, and another study indicated that compared to traditional visual methods, CAD/CAM-based tools showed higher assessment precision in evaluating taper features in crown preparation (Mays et al., 2016). Another study assessed the inter- and intra-grader agreements of five preclinical instructors using PrepCheck Sirona for crown preparation (Miyazono et al., 2019). The results showed that intra-grader agreement in PrepCheck was significantly higher than that achieved using traditional visual grading methods and that inter-grader disagreement presented the opposite pattern. Romexis Compare 2.0 (Compare) and prepCheck 1.1 (Sirona) systems were used to assess crown preparations digitally, and these findings were compared to those of traditional faculty assessments (Uoshima et al., 2021). Conventional instructions were provided to the students prior to the assessment of crown preparations. Technical scores did not differ between student groups for any of the assessment approaches, and a student survey reported that the integration of digital tools was preferred. Digital assessment of a tooth prepared by students significantly reduces variability and subjectivity compared with traditional visual inspection (Kateeb et al., 2017). Although digital assessment can overcome some limitations, it cannot completely replace traditional assessment methods (Seet et al., 2020).

Another study found no significant difference between the two grading systems evaluating 98 Class I preparations, with the benefit of software as a self-assessment tool for students wanting to improve their skills (Sly et al., 2017). Medit Compare, a digital feedback tool, was compared with the traditional putty index in a study comparing satisfaction scores and the educational impact of tooth preparation. The results showed no significant differences; however, a high level of 2023 was achieved using digital tools, enhancing educational satisfaction and allowing for an objective evaluation (Kim et al., 2023). A recent systematic review showed that using computer simulation to teach skills related to restorative dentistry to undergraduates allows simulation exercises that are both beneficial and effective (Nassar and Tekian, 2020). However, some students complained about not having enough scanners or time to practice, while others thought that the policy was too strict (Schepke et al., 2020).

To the best of our knowledge, no previous study has evaluated the intra- and inter-grade reliability of assessments for preclinical Class II cavity preparation for resin composite restorations or compared the outcomes between digital and visual evaluation methods. Therefore, this study aimed to evaluate the intra- and inter-grader reliabilities of four evaluators using three different digital intraoral scanners and visual methods for typodontic Class II composite preparations. The null hypothesis was that there is no difference in intra- and inter-examiner reliability between visual grading method (VGM) and digital grading method (DGM) methods and that there is no difference in inter-examiner reliability between different evaluators in the evaluation of Class II composite preparations.

2. Materials and methods

2.1. Sample selection

The King Abdulaziz University Institutional Review Board (#150–12-22) approved this study for scientific and educational research. The sample size was determined based on previous studies, adding a 5 % marginal error (N = 95) using the following equation: [(98 + 100 + 74) /3 + 5 % = 95]. Eventually, 95 of the 120 typical teeth were randomly selected from the bowl. The collected teeth were Class II mesial composite preparations of the mandibular second molars with adjacent teeth. Any other cavity preparations for composite restorations or composite preparations of either the maxillary molars or teeth without adjacent teeth, or both, were excluded. Third-year dental students were given a lecture and demonstration video on preparing a Class II cavity for composite restoration using specific criteria, following the King Abdulaziz University Faculty of Dentistry (KAUFD) rubric (Table 1). A total of 120 dental students prepared the mandibular second molars. The teeth were collected anonymously and identified by numbers.

2.2. Grading methods

Four calibrated restorative dentist examiners (E1-E4), two academicians (E1, E3), and two clinicians (E2, E4) assessed the 95 typodont tooth preparations twice according to the modified KAUFD rubric, once by the traditional visual (VGM) and once by digital (DGM) grading methods, using three types of IOS: Trios 4 (3Shape, Denmark), i700 (Medit, Korea), and Shining 3D Aoralscan 3 (Shinning 3D, China) that use Medit Link software (Medit Design, Korea). Both restorative dentists and academics have strong expertise in clinical judgment based on research and practice. Calibration training was performed by demonstrating how to evaluate cavity preparation using the Medit Design application and measuring tools in the Medit Link software digital expert, and each one was individually tested until they reached 100 % accuracy. The scanning process was performed by a prosthodontist, who was not among the examiners, by scanning the prepped typodontist's teeth and storing the scans on a computer for later examination by four evaluators. Google Forms, with a specific set of questions, were created for data entry, either from visual or digital evaluations (Fig. 1), and each examiner was given a specific ID (E1-E4). The order of the evaluations was random for both methods.

2.2.1. Visual grading method (VGM)

The evaluators took a random tooth from a bowl, held it in their hands, and assessed it using a UNC-15 periodontal probe. The seven criteria of the KAUFD rubric (Table 1) included cavity width, extension, pulpal and gingival depths, pulpal and gingival floor uniformity, buccal and lingual wall directions, and finishing level of the walls are all checked.

2.2.2. Digital grading method (DGM)

For digital evaluation, scans were evaluated after randomization. All scans were gathered together and organized according to the file size. Ten scans were taken from the top of the list of scans, 10 from the bottom, and 10 from the middle were selected first and then the process is repeated to reduce the risk of bias. The seven criteria were modified from the KAUFD rubric (Table 1) and included cavity width, extension, pulpal and gingival depths, pulpal and gingival floor uniformity, buccal and lingual wall directions, and finishing level of the walls. The scores for each item are 0, 1, or 2 (Table 1). A total of 285 scans (n = 95 from each IOS group) were imported randomly into the Medit Link software (Medit Design, Korea) in standard tesselation language (STL) format, where each scan was named with two digits: the number of teeth and the letter that symbolized the scanner type (A: Medit i700, B: Trios 4, C: Shining). To ensure blinding and bias control, all examiners first

Table 1
King Abdulaziz University Faculty of Dentistry (KAUFD) rubric for class II cavity preparation for composite restoration.

Stage	Principle	Grade 2	1	0
Outline form	Width	Buccolingual width = 1/4 the intercuspal distance	Buccolingual width >1/4 < 1/2 the intercuspal distance	Buccolingual width < 1/4 or > 2/3 the intercuspal distance
	Extension	Proper Primary pits and fissures are included AND 2 mm thickness of the remaining marginal ridge	Under-extended Does not include all the primary pits and fissures OR > 2 mm thickness of the remaining marginal ridge	Over-extended Extends to the cusps and marginal ridge OR < 2 mm thickness of the remaining marginal ridge
	Pulpal depth	Proper and even 1.5 to 2.5	Uneven cavity depth but within the average limit	Shallow or deep <1.5 or > 2.5
	Gingival depth	0.5–1 mm	>1 but ≤ 1.5	<0.5 or > 1.5
Primary resistance form	Pulpal and gingival floor	Uniform (no irregularities in both pulpal and gingival floor)	Half-uniform (irregularities in one floor, either pulpal or gingival floor)	Not uniform (irregularities in both pulpal and gingival floor)
	Wall direction Proximal	Converged (both buccal and lingual wall ≤ 90 degrees to the floor)	Half converged (one wall, either buccal or lingual wall ≤ 90 degrees to the floor)	Diverged (Both walls > 90 degrees to the floor)
Finishing of walls	Walls	Smooth (Roughness less than 20 %)	Half-smooth (Roughness 20–50 %)	Rough (Roughness more than 50 %)

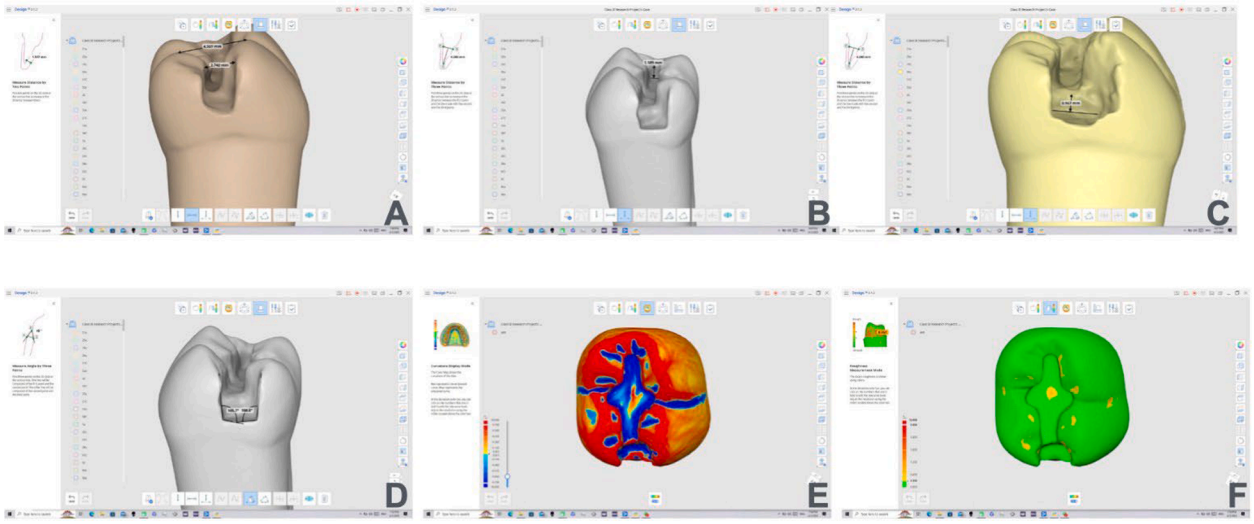

Fig. 1. Comprehensive diagrams illustrating various measurement techniques for Class II cavity preparation using Medit Design App in Medit Link software: (A) Width measurement using a two-point distance tool; (B) Pulpal depth measurement at the distal wall using a three-point distance tool; (C) Gingival depth measurement from the pulpal to the gingival floor with a three-point distance tool; (D) Wall direction measurement indicating buccal and lingual angles with a three-point angle tool; (E) Pulpal and gingival floor uniformity displayed in color-coded curvature mode, highlighting areas of un-uniformity; (F) Surface roughness assessment using a roughness measurement mode, with rough areas indicated in yellow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

assessed the teeth using VGM, followed by DGM. The four examiners were trained using software and a rubric (Table 1), and the scores were calibrated individually. The second evaluation for inter-reliability was performed after a period of one week.

2.2.3. Evaluation rubric

The rubric used in the current study consisted of three stages (outline form, primary resistance form, and finishing of walls) and seven principles of tooth preparation, graded on a scale of 0–2 points (Table 1). The outline form was evaluated on the basis of four principles: intercuspal width, cavity extension, and pulpal and gingival depths of the cavity. The primary form of resistance was evaluated based on the uniformity of the pulpal and gingival floors and the direction of the proximal walls. Finally, the finishing and smoothness of the walls are evaluated. The grading in the current rubric is based on the scoring of the unfulfilled criteria for each assessment item (Fig. 1).

2.3. Statistical analysis

All statistical analyses were performed using SPSS version 26 software (IBM Corp., Armonk, New York, USA). The values represent means ± standard deviations for continuous variables and counts and percentages for categorical variables. Two-way ANOVA was used to assess inter- and intra-grader agreements in the digital assessment tools, which were then compared with the visual assessment tool. Pairwise comparisons were performed for each simple main effect, with 95 % confidence intervals and Bonferroni-adjusted p-values. Statistical significance was set at $p < 0.05$ for two-way interactions and main effects.

A two-way analysis of variance (ANOVA) was conducted to examine the effects of the examiner and evaluation methods (VGM vs. DGM) on students' total scores for Class II composite cavity preparation. Residual analysis was performed to test the assumptions of the two-way ANOVA. Outliers were assessed by inspecting boxplots. Normality of distribution was assessed using the Shapiro–Wilk normality test for each cell of the design, and homogeneity of variance was assessed using the Levene test.

3. Results

3.1. Descriptive analysis

Table 2 shows the mean and standard deviation for each examiner and evaluation technique.

3.2. Comparative analysis

No outliers were present, and the residuals exhibited a normal distribution ($P > 0.05$). Homogeneity of variance was confirmed ($P = 0.061$). A statistically significant interaction emerged between examiners and the evaluation technique employed to assess the total score of the prepped teeth: $F(9, 1504) = 3.893$, $P = 0.001$, partial $\eta 2 = 0.023$ (Table 3). Therefore, an analysis of simple main effects for the evaluation methods was performed using Bonferroni adjustment and was accepted at $P < 0.025$. The results showed a statistically significant difference in the mean total score between examiners using VGM: $F(3, 1504) = 2.705$, $P = 0.034$, partial $\eta 2 = 0.005$ (Table 3). For VGM, the mean total score for E1 was 8.55 (SD = 2.364) and 7.76 (SD = 2.474) points for E2, 8.00 (SD = 2.961) points for E3, and 8.17 (SD = 2.738) points for E4 (Table 3). This finding indicates higher inter-grade reliability when using the DGM than when using the VGM.

In addition, a significantly higher intra-grade consistency was found when using the DGM than when using the VGM. (Figs. 2 and 3), respectively. Pairwise comparisons were performed for each simple main effect, with 95 % confidence intervals and Bonferroni-adjusted P values. A significant difference in the mean total score per examiner (intragrader) was observed between the VGM and DGM groups (Table 3). A significantly lower ($P < 0.05$) intra-grader consistency in the final score was observed when using VGM. However, when the DGM was used, the intra-grade consistency of the final scores improved. These findings suggest that the DGM may improve assessment reliability compared to the VGM.

4. Discussion

The results revealed a statistically significant interaction between examiners and evaluation techniques used in this context. In particular, a higher inter-grade agreement was found when using DGM than when using VGM. In addition, significantly higher intragrade reliability was observed when grading with the DGM than with the VGM. These findings suggest that using the DGM may lead to more reliable grading, either between for the same evaluator or between evaluators.

To the best of our knowledge, this study is the first to provide a novel digital evaluation protocol based on a validated and tested rubric to evaluate student performance in Class II cavity preparation for composites. As we all know, studies evaluating DGM and VGM in dental training are scarce. DGM tools may facilitate students' educational experiences and allow instructors to provide more objective and consistent feedback. This study evaluated the differences between digital and visual evaluation techniques for Class II cavity preparation using three different IOS and four examiners. The grades were more consistent when using IOS-based DGM than when using VGM. Previous studies reported similar findings (Nagy et al., 2018; Greany et al., 2019; Jorquera et al., 2021; Kim et al., 2023). None of these studies evaluated Class II cavity preparation for Class II composites. Instead, they focused on evaluating wax-ups digitally or in crown preparations.

Technical skills, knowledge, and attitudes are components of dental practice. Technical skills, including cognitive, psychomotor, and effective skills, must be objectively assessed for high validity and fairness. Comprehensive, frequent, accurate, and consistent feedback improves students' learning experience. Digital 3D scanned models have revolutionized dental education by providing students with detailed visualization and analysis of prepped teeth. They also provide immediate feedback, allowing students to quickly understand and correct their mistakes, enabling the repeatability of procedures, and promoting a standardized form of training. They also enable students to self-assess their work and promote reflective practice. However, challenges, such as accessibility and cost, learning curves, and integration with traditional methods, still exist. Therefore, the integration of digital and conventional training approaches should be carefully planned to ensure that all students benefit equally from both methods. A combination of various technical skill assessment methods can improve validity (Lee et al., 2018). The different methods of technical skills training include 1) conservative skills training with typodonts, 2) simulation models, 3) the use of virtual reality technology, and 4) augmented reality technology. Furthermore, multiple technical skill assessment methods have been described, including those for clinical skills and those for clinical skills, which may include rubrics, global rating scales, structured rating scales, and testing typodonts.

The uptake of DGM is increasing and may soon become a standard in educational settings. This shift may have helped instructors provide more accurate feedback. In our study, inter- and intra-grader agreements were assessed for DGM and VGM when evaluating Class II cavities. The DGM had higher intra- and inter-grade agreements. Several other studies have reported similar patterns when assessing crown preparations (Gratton et al., 2015; Gratton et al., 2016). Additionally, students tend to be more satisfied when digital assessment tools are used because they provide more objective evaluations (Clancy et al., 2002).

Not all assessment tools stand-alone, and practical feedback from supervisors is crucial. A direct explanation of the mistake, why it happened, how it can be prevented/corrected, and giving students frequent constructive feedback are essential, especially in preclinical courses, where students have little to no experience. Combining these efforts with a digital assessment workflow can enhance students' learning experience. Students may be able to zoom in and out of their work and attempt to assess it. However, a major drawback of DGM is the lack of live feedback during preparation, the need for specialized equipment, and steep user learning curves (Gratton et al., 2016).

"Programmatic assessment" is a novel method that integrates multiple assessment methods to ensure effective learning and increase the reliability and validity of educational assessments (Mays et al., 2016). This concept should be explored further in future studies. In this study, we combined a novel digital protocol with an institutionally validated evaluation rubric. Our DGM protocol for Class II cavity preparation for

Table 2
Dependent Variable: Total Score.

Examiner ID	Evaluation Technique	Mean	Std. Deviation	N
E1	Digital Medit	6.47	1.636	95
	Digital Trios	6.15	1.444	95
	Digital Shining	6.27	1.691	95
	Visual	8.55	2.364	95
	Total	6.86	2.060	380
E2	Digital Medit	6.16	1.639	95
	Digital Trios	5.19	1.240	95
	Digital Shining	5.63	1.495	95
	Visual	7.76	2.474	95
	Total	6.18	2.016	380
E3	Digital Medit	7.01	1.954	95
	Digital Trios	6.72	1.950	95
	Digital Shining	7.23	2.013	95
	Visual	8.00	2.961	95
	Total	7.24	2.301	380
E4	Digital Medit	6.45	1.514	95
	Digital Trios	6.75	1.598	95
	Digital Shining	7.18	1.676	95
	Visual	8.07	2.738	95
	Total	7.11	2.033	380
Total	Digital Medit	6.52	1.715	380
	Digital Trios	6.20	1.695	380
	Digital Shining	6.58	1.847	380
	Visual	8.09	2.650	380
	Total	6.85	2.143	1520

Table 3
Pairwise comparisons.

Pairwise Comparisons Dependent Variable: Total Score							
Examiner ID	(I) Evaluation Technique	(J) Evaluation Technique	Mean Difference (I-J)	Std. Error	Sig.[b]	95 % Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
E1	Digital Medit	Digital Trios	0.326	0.284	1.000	−0.424	1.077
		Digital Shining	0.200	0.284	1.000	−0.551	0.951
		Visual	−2.074*	0.284	0<.001	−2.824	−1.323
	Digital Trios	Digital Medit	−0.326	0.284	1.000	−1.077	0.424
		Digital Shining	−0.126	0.284	1.000	−0.877	0.624
		Visual	−2.400*	0.284	0<.001	−3.151	−1.649
	Digital Shining	Digital Medit	−0.200	0.284	1.000	−0.951	0.551
		Digital Trios	0.126	0.284	1.000	−0.624	0.877
		Visual	−2.274*	0.284	0<.001	−3.024	−1.523
	Visual	Digital Medit	2.074*	0.284	0<.001	1.323	2.824
		Digital Trios	2.400*	0.284	0<.001	1.649	3.151
		Digital Shining	2.274*	0.284	0<.001	1.523	3.024
E2	Digital Medit	Digital Trios	0.968*	0.284	0.004	0.218	1.719
		Digital Shining	0.526	0.284	0.385	−0.224	1.277
		Visual	−1.600*	0.284	0<.001	−2.351	−0.849
	Digital Trios	Digital Medit	−0.968*	0.284	0.004	−1.719	−0.218
		Digital Shining	−0.442	0.284	0.720	−1.193	0.309
		Visual	−2.568*	0.284	0<.001	−3.319	−1.818
	Digital Shining	Digital Medit	−0.526	0.284	0.385	−1.277	0.224
		Digital Trios	0.442	0.284	0.720	−0.309	1.193
		Visual	−2.126*	0.284	0<.001	−2.877	−1.376
	Visual	Digital Medit	1.600*	0.284	0<.001	0.849	2.351
		Digital Trios	2.568*	0.284	0<.001	1.818	3.319
		Digital Shining	2.126*	0.284	0<.001	1.376	2.877
E3	Digital Medit	Digital Trios	0.295	0.284	1.000	−0.456	1.046
		Digital Shining	−0.221	0.284	1.000	−0.972	0.530
		Visual	−0.989*	0.284	0.003	−1.740	−0.239
	Digital Trios	Digital Medit	−0.295	0.284	1.000	−1.046	0.456
		Digital Shining	−0.516	0.284	0.418	−1.267	0.235
		Visual	−1.284*	0.284	0<.001	−2.035	−0.533
	Digital Shining	Digital Medit	0.221	0.284	1.000	−0.530	0.972
		Digital Trios	0.516	0.284	0.418	−0.235	1.267
		Visual	−0.768*	0.284	0.042	−1.519	−0.018
	Visual	Digital Medit	0.989*	0.284	0.003	0.239	1.740
		Digital Trios	1.284*	0.284	0<.001	0.533	2.035
		Digital Shining	0.768*	0.284	0.042	0.018	1.519
E4	Digital Medit	Digital Trios	−0.295	0.284	1.000	−1.046	0.456
		Digital Shining	−0.726	0.284	0.064	−1.477	0.024
		Visual	−1.621*	0.284	0<.001	−2.372	−0.870
	Digital Trios	Digital Medit	0.295	0.284	1.000	−0.456	1.046
		Digital Shining	−0.432	0.284	0.774	−1.182	0.319
		Visual	−1.326*	0.284	0<.001	−2.077	−0.576
	Digital Shining	Digital Medit	0.726	0.284	0.064	−0.024	1.477
		Digital Trios	0.432	0.284	0.774	−0.319	1.182
		Visual	−0.895*	0.284	0.010	−1.646	−0.144
	Visual	Digital Medit	1.621*	0.284	0<.001	0.870	2.372
		Digital Trios	1.326*	0.284	0<.001	0.576	2.077
		Digital Shining	0.895*	0.284	0.010	0.144	1.646

Based on estimated marginal means
*. The mean difference is significant at the 0.05 level.
b. Adjustment for multiple comparisons: Bonferroni.

composite fillings offers a more objective and precise alternative to VGM, ensuring consistency and reliability of student evaluations. DGM uses predefined criteria, reduces evaluator bias, and provides detailed feedback for learning. This approach also enhances the implementation of programmatic assessments in dental education, allowing for consistent longitudinal assessments, improved inter-rater reliability, and data-driven decision-making. Implementing our DGM protocol allows for accurate measurements of student progress, ensuring that improvements or regressions in skills are captured accurately. It also improves inter-grader reliability, ensuring that assessments are comparable regardless of the evaluator. This aligns with the principles of programmatic assessment, which enhance the accuracy of performance tracking and the quality of feedback provided to students. This principle supports a more structured and data-driven approach to dental education, ultimately leading to better professional dental preparation.

Future studies should also compare student performance with technical skill self-assessment. Teachers or instructors can set a particular tolerance level for either the required amount of tooth removal preparation or both. Meanwhile, a color-coded scale can show the students instantly if they are under or overprepared and if the wall direction angle exceeds or is insufficient compared to the data in the source file. This approach can encourage student self-assessment and provide on-demand feedback outside regular training hours, thereby supporting skill development. The main limitation of this study was that it was performed in a preclinical setting on only one tooth: the mandibular second molar. Second, cost and convenience play major roles in limiting the implementation of the tested digital methods for assessment.

5. Conclusion

The DGM showed higher reliability within and between evaluators than the VGM. Digital assessment tools can provide students with

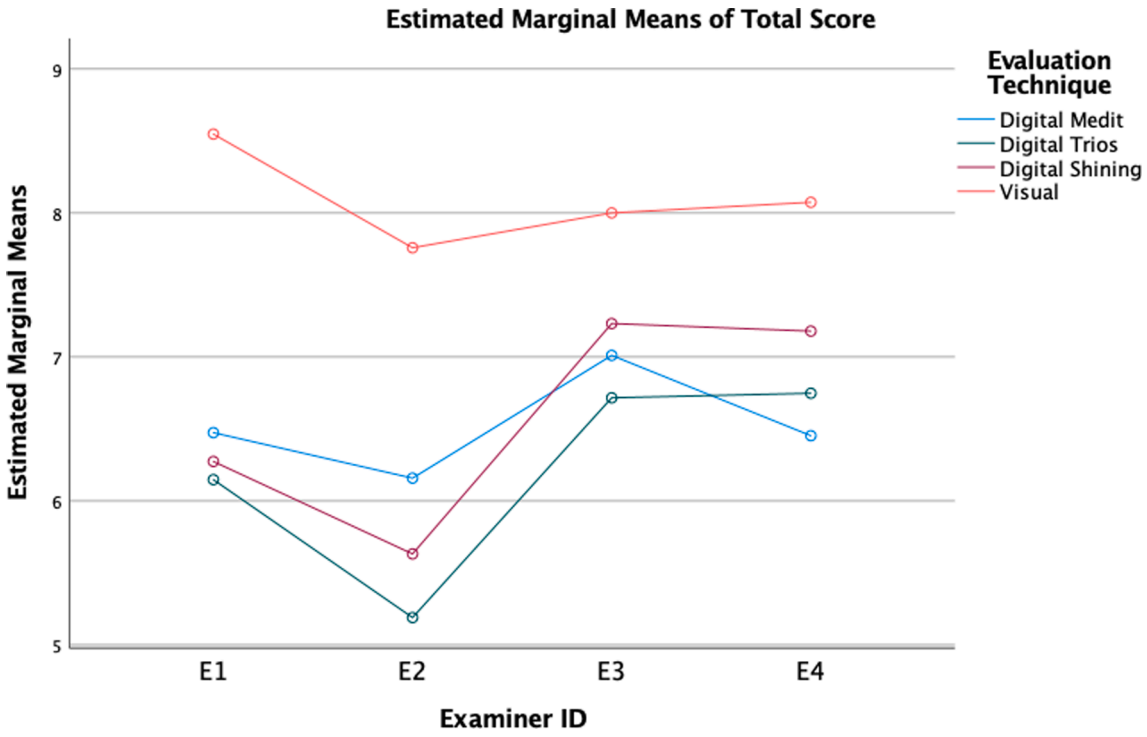

Fig. 2. Estimated marginal means of total score based on the evaluation technique.

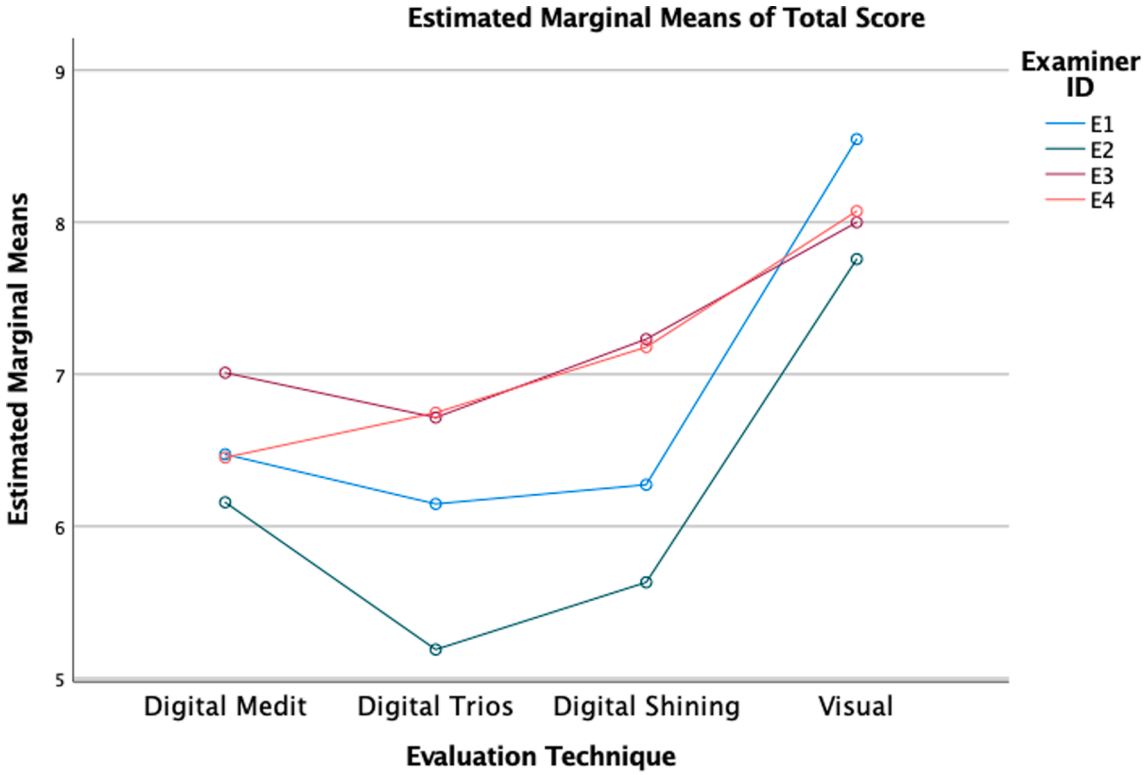

Fig. 3. The estimated marginal mean of the total score is based on the evaluator.

consistent and objective feedback when preparing Class II composites. Overall, the present study highlights the potential benefits of using DGM in dental education and underscores the importance of considering the effects of the examiner and evaluation methods on student performance metrics.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Amr Ahmed Azhari: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Walaa Magdy Ahmed:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Raghad Almaghrabi:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Aroob Almalki:** Data curation, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Yasser Merdad:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft. **Hind Alrefai:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft. **Mawadah T. Bahrawi:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft. **Lama Faisal Kordi:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Authors would like to thank the internship training program at the faculty of dentistry, King Abdulaziz University, Jeddah, Saudi Arabia for their support throughout the research project.

References

Clancy, J.M.S., Lindquist, T.J., Palik, J.F., Johnson, L.A., 2002. A comparison of student performance in a simulation clinic and a traditional laboratory environment: Three-year results. J. Dent. Educ. 66, 1331–1337. https://doi.org/10.1002/j.0022-0337.2002.66.12.tb03606.x.

Gratton, D.G., Kwon, S.R., Blanchette, D., Aquilino, S.A., 2016. Impact of digital tooth preparation evaluation technology on preclinical dental students' technical and self-evaluation skills. J. Dent. Educ. 80, 91–99. https://doi.org/10.1002/j.0022-0337.2016.80.1.tb06062.x.

Gratton, D.G., Kwon, S.R., Blanchette, D.R., Aquilino, S.A., 2017. Performance of two different digital evaluation systems used for assessing preclinical dental students' prosthodontic technical skills. Eur. J. Dent. Educ. 21, 252–260. https://doi.org/10.1111/eje.12231.

Greany, T.J., Yassin, A., Lewis, K.C., 2019. Developing an all-digital workflow for dental skills assessment: Part I, visual inspection exhibits low precision and accuracy. J. Dent. Educ. 83, 1304–1313. https://doi.org/10.21815/JDE.019.132.

Jorquera, G., Sánchez, J.P., Sampaio, C.S., Atria, P., Fernández, E., 2021. Improvement preclinical and clinical skills for dental preparations using assisted training software. Eur. J. Dent. Educ. 25, 856–863. https://doi.org/10.1111/eje.12665.

Kateeb, E.T., Kamal, M.S., Kadamani, A.M., Abu Hantash, R.O., Abu Arqoub, M.M., 2017. Utilising an innovative digital software to grade preclinical crown preparation exercise. Eur. J. Dent. Educ. 21, 220–227. https://doi.org/10.1111/eje.12204.

Kim, Y.K., Kim, J.H., Jeong, Y., Yun, M.J., Lee, H., 2023. Comparison of digital and conventional assessment methods for a single tooth preparation and educational satisfaction. Eur. J. Dent. Educ. 27, 262–270. https://doi.org/10.1111/eje.12799.

Lam, M.T., Kwon, S.R., Qian, F., Denehy, G.E., 2015. Evaluation of an innovative digital assessment tool in dental anatomy. J. Contemp. Dent. Pract. 16, 366–371. https://doi.org/10.5005/jp-journals-10024-1691.

Lee, C., Kobayashi, H., Lee, S.R., Ohyama, H., 2018. The role of digital 3D scanned models in dental students' self-assessments in preclinical operative dentistry. J. Dent. Educ. 82, 399–405. https://doi.org/10.21815/JDE.018.046.

Mays, K.A., Crisp, H.A., Vos, P., 2016. Utilizing CAD/CAM to measure total occlusal convergence of preclinical dental students' crown preparations. J. Dent. Educ. 80, 100–107. https://doi.org/10.1002/j.0022-0337.2016.80.1.tb06063.x.

Miyazono, S., Shinozaki, Y., Sato, H., Isshi, K., Yamashita, J., 2019. Use of digital technology to improve objective and reliable assessment in dental student simulation laboratories. J. Dent. Educ. 83, 1224–1232. https://doi.org/10.21815/JDE.019.114.

Nagy, Z.A., Simon, B., Tóth, Z., Vág, J., 2018. Evaluating the efficiency of the dental teacher system as a digital preclinical teaching tool. Eur. J. Dent. Educ. 22, e619–e623. https://doi.org/10.1111/eje.12365.

Nassar, H.M., Tekian, A., 2020. Computer simulation and virtual reality in undergraduate operative and restorative dental education: A critical review. J. Dent. Educ. 84, 812–829. https://doi.org/10.1002/jdd.12138.

Renne, W.G., McGill, S.T., Mennito, A.S., Wolf, B.J., Marlow, N.M., Shaftman, S., Holmes, J.R., 2013. E4D compare software: An alternative to faculty grading in dental education. J. Dent. Educ. 77, 168–175. https://doi.org/10.1002/j.0022-0337.2013.77.2.tb05459.x.

Schepke, U., van Wulfften Palthe, M.E., Meisberger, E.W., Kerdijk, W., Cune, M.S., Blok, B., 2020. Digital assessment of a retentive full crown preparation-An evaluation of prepCheck in an undergraduate preclinical teaching environment. Eur. J. Dent. Educ. 24, 407–424. https://doi.org/10.1111/eje.12516.

Seet, R.H., Soo, P.R., Leong, K.J.M., Pang, J.J.H., Lee, F.K.F., Tan, M.Y., 2020. Crown preparations by undergraduate dental students: A comparison of conventional versus digital assessment via an intraoral scanner. J. Dent. Educ. 84, 1303–1313. https://doi.org/10.1002/jdd.12285.

Sly, M.M., Barros, J.A., Streckfus, C.F., Arriaga, D.M., Patel, S.A., 2017. Grading Class I preparations in preclinical dental education: E4D compare software vs. the traditional standard. J. Dent. Educ. 81, 1457–1462. https://doi.org/10.21815/JDE.017.107.

Stoilov, M., Trebess, L., Klemmer, M., Stark, H., Enkling, N., Kraus, D., 2021. Comparison of digital self-assessment systems and faculty feedback for tooth preparation in a preclinical simulation. Int. J. Environ. Res. Public Health. 18, 13218. https://doi.org/10.3390/ijerph182413218.

Tiu, J., Cheng, E., Hung, T.C., Yu, C.C., Lin, T., Schwass, D., Al-Amleh, B., 2016. Effectiveness of crown preparation assessment software as an educational tool in simulation clinic: A pilot study. J. Dent. Educ. 80, 1004–1011. https://doi.org/10.1002/j.0022-0337.2016.80.8.tb06181.x.

Uoshima, K., Akiba, N., Nagasawa, M., 2021. Technical skill training and assessment in dental education. Jpn. Dent. Sci. Rev. 57, 160–163. https://doi.org/10.1016/j.jdsr.2021.08.004.