RESEARCH ARTICLE

Utilizing 3D Model and Comparison Software Tools in Improving Assessment Methods: A Pilot Study

Michelle Aguilos Thompson[1,*], Marilia Mattos Sly[2], Maria de Lourdes Gonzalez[2], and Shalizeh Azizi Patel[3]

ABSTRACT

Purpose/Objectives: **This manuscript investigates the effectiveness of computer-assisted design and manufacturing (CAD/CAM) technology in preclinical dental education. The COVID-19 pandemic created a need for remote teaching approaches. In preclinical fixed prosthodontics courses, traditional evaluation methods rely on faculty to visually grade laboratory projects utilizing rubrics and 2D images. Assessment and feedback vary amongst faculty, and faculty calibration has been inconsistent in reducing intra-and inter-rater variability. This study aimed to compare faculty assessment of student full coverage crown (FCC) preparations traditionally versus digitally through 3D model visual assessment and/or comparative analysis software reports.**

Methods: **Eighteen UTHealth Houston School of Dentistry faculty participated, each randomly assigned into one of three assessment groups: traditional visual Group (TV), 3D Group (3D), and 3D visual assessment supplemented with comparative analysis software report (3D+) Group. Faculty graded the same eight lithium disilicate typodont preparations for all three assessment groups.**

Results: **Survey results showed the ease of use and efficiency among the three assessment methods varied significantly. Grading scores revealed that axial wall reduction significantly differed among the three assessment methods.**

Conclusion: **This study concludes that the combination of the two digital assessment methods can overcome the limitations of traditional evaluation methods and digital methods alone by providing objective feedback to students and enhancing their learning experience.**

Keywords: Dental education, digital dentistry, educational assessment, self-assessment.

1. INTRODUCTION

The in-person teaching approach widely practiced during the simulation and preclinical sessions became a challenge during the COVID-19 pandemic. Dental educators modified their teaching practices and used emerging technologies to teach while social distancing. These reactive teaching methodologies were valuable to student learning at the time, but further exploration is required to evaluate their effectiveness in the current landscape as students and teachers return to classroom and simulation laboratories. Any future teaching approach, remote or in-person, must encourage the application of knowledge, improve hand-motor skills, inspire dialogue, and provide autonomy for both the learner and the educator.

In preclinical fixed prosthodontics courses, traditional evaluation methods rely on faculty to visually grade simulation exercises utilizing rubrics and 2D images, followed by verbal and/or written feedback. However, assessment and feedback vary amongst faculty, which can be discouraging and impede student learning [1], [2]. Faculty calibration has failed to reduce intra and inter-rater variability. In addition, increased class sizes and time constraints due to an already-packed curriculum pose limitations on faculty's one-on-one interactions with students in providing meaningful feedback [3]. Thus, creating a practical model to overcome these challenges may be the

next step. Computerized evaluation of student projects has been proposed to serve as an objective tool [4].

Computer-assisted design and manufacturing (CAD/-CAM) technology, including intraoral scanning, has been implemented in predoctoral education for the design and fabrication of indirect dental restorations and prostheses. Comparative analysis software associated with this technology, such as PrepCheck® and Romexis Compare®, is utilized for objective evaluation of students' projects. With CAD/CAM, students are able to practice exercises independently by digitally scanning their preparation, comparing it to an ideal model, and receiving an objective software report on their work. This software has shown promise as a self-assessment tool for students [5]. Further, it has improved intra-and inter-rater agreement in the grading of student preparations in student simulation laboratories [6]–[9]. It was found to be more reliable and precise than traditional grading by reducing subjectivity [10]. However, students felt that the software was not a replacement for faculty as it cannot grade all aspects of a crown preparation such as smoothness of the finish line [11].

Comparative analysis software alone cannot replace faculty visual grading of preparations. Visual grading of 3D scanned models by faculty has been shown to be a valid tool. Lee *et al*. evaluated the use of 3D scanned models for student self-assessment and faculty visual grading [12]. Students and faculty evaluated 3D scanned models, using zoom, pan, and rotation functions. Faculty scores of operative preparations showed no difference between conventional and digital visual grading methods, thus validating this digital assessment tool. Criteria including occlusal clearance, smoothness of preparation finish line and other features cannot be measured by comparative analysis software, so a visual assessment by faculty of 3D models can bridge this gap, especially in a remote environment.

A combination of comparative analysis software-generated reports with 3D model visual grading by faculty has yet to be explored. These digital methods can be applied by faculty in person or remotely for grading of student preparations. Previous studies have utilized cameras to capture 2D images of student preparations for faculty grading remotely. However, 3D evaluation in a remote environment needs further evaluation [13]. To address both remote learning and to increase the objectivity of grading, the investigators of this study proposed developing an objective and socially distanced feedback mechanism for grading students' psychomotor skills. This study aimed to compare faculty grading of student full coverage crown (FCC) preparations traditionally versus digitally through 3D model visual assessment and/or comparative analysis software reports. The null hypothesis was there would be no difference in faculty grading of FCC using traditional visual assessment, 3D models visual assessment, and/or comparative analysis software reports.

2. Materials and Methods

The study protocol was approved by the Committee for the Protection of Human Subjects (CPHS) of UTHealth Houston, HSC-DB-20-1284. Second-year dental students completed lithium disilicate crown preparations on a lower right first molar (#30) using a Kilgore Series 200 typodont (Coldwater, MI, USA), as a simulation exercise for a Digital Dentistry preclinical course. Eighteen faculty volunteered to participate in this study and were randomly assigned into one of three assessment groups: (1) Traditional visual, Group TV, (2) 3D model, Group 3D, and (3) 3D model supplemented with comparative analysis software report, Group 3D+. For all three assessment methods, faculty graded eight lithium disilicate typodont preparations. A standard rubric was used to evaluate the preparations and the following criteria were assessed: (1) occlusal reduction, (2) axial wall reduction, (3) axial wall taper, (4) marginal width, (5) cervical finish line level with respect to gingiva, and (6) margin proximity to adjacent teeth (Table I).

A calibration session was provided to faculty participants on grading rubrics and workflows. The faculty graded independently and the investigators were only available for instructional technology support. Faculty participants were asked to submit grades on Qualtrics® for each preparation (Provo, UT, USA); the rubric was presented on a 5-point Likert scale. Following grading, faculty were asked to complete a survey via Qualtrics® on their perception of the usefulness of their assigned grading method and the feasibility for its future integration into the preclinical prosthodontics curriculum (Table II). Faculty survey responses included (1) strongly agree, (2) agree, (3) neutral, (4) disagree, (5) strongly disagree.

2.1. Assessment Methods

2.1.1. Traditional Visual Group (TV)

Seven faculty assigned to group TV visually assessed student preparations on a typodont. A mirror, explorer, periodontal probe, and eye magnification (loupe × 2.5) were utilized.

2.1.2. 3D Model Group (3D)

Six faculty assigned to this group were given access to pre-scanned student preparations. Planmeca Emerald® scanners and Romexis® software (Hoffman Estates, IL, USA) were utilized by the investigators to generate 3D model scans, visible on a computer monitor.

2.1.3. 3D Model Supplemented with Comparative Analysis Software Report Group (3D+)

Six faculty assigned to this group were provided with the same instructions as group 3D and also given access to Romexis Compare® software (Hoffman Estates, IL, USA) to digitally compare each student's preparation with the ideal Master Preparation (MP) illustrated in Fig. 1. The tolerance was set to 350 μm as faculty assessed differences between each student's preparation and the MP.

2.2. Statistical Analysis

Kruskal Wallis tests (KW) were performed to evaluate if differences occurred among the three methods. Tukey-Kramer-Nemenyi multiple comparison tests were performed to assess pairwise differences among the three groups when the KW test was significant. An ANOVA

TABLE I: Ideal Preparation Guidelines

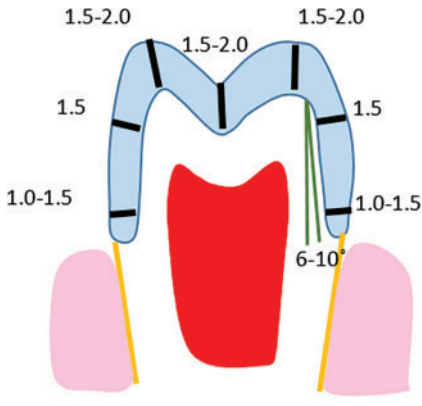| Evaluated criteria	Preparation guidelines
Occlusal reduction	1.5–2.0 mm
Axial wall reduction	1.5 mm
Axial wall taper	6–10 degrees
Marginal width	1.0–1.5 mm
Margin level	Supragingival
Margin level to adjacent tooth	1.0 mm

TABLE II: Faculty Survey Questions

| Faculty survey questions |
|---|
| The assessment method was easy to use. |
| I had adequate time to use the assessment method. |
| The assessment method enables providing constructive feedback to student learners. |
| The assessment method can be used for faculty calibration. |
| The assessment method can be used for objective feedback to students in a virtual meeting. |
| Open box comments: Please share your thoughts and comments on utilizing the assessment method and its future use in the preclinical curriculum. |
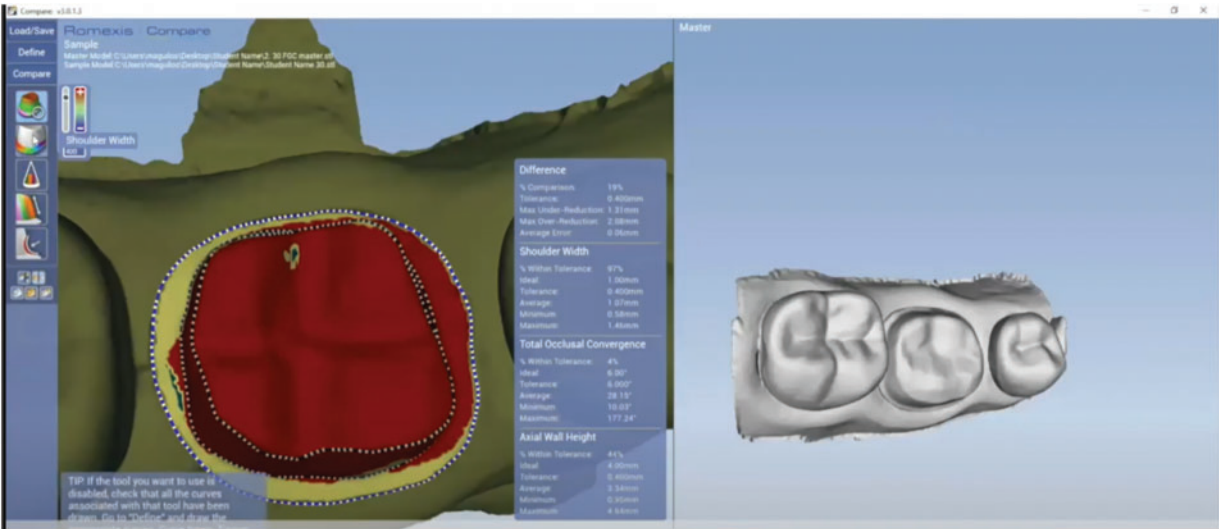
Fig. 1. Comparison of tooth #30 student preparation and MP using Romexis Compare®.

was used to statistically analyze the mean of each variable amongst the three methods. All statistical analyses were performed using R statistical software with a $p < 0.05$ indicating significant differences [14].

3. Results

Table III illustrates the descriptive statistics results including chi-square, *p*-value, mean value, and pairwise comparison for Faculty Survey. KW indicated that the ease of use among the three assessment methods varied *significantly*, driven mainly by the difference between 3D+ and TV groups. Graders in the 3D+ Group expressed that the assessment method was more difficult to use compared to the TV group. The KW showed that efficiency among the three assessment methods also varied *significantly*, marginally driven by the difference between 3D+ and 3D groups. The 3D Group was considered marginally more efficient than the 3D+ Group. All three groups of evaluators stated that each method is useful for faculty calibration and groups 3D and 3D+ agreed that it provided objective feedback in a virtual setting. Responses were neutral for the TV Group in regards to providing objective feedback in a virtual setting. Overall, there was no difference among the three assessment methods on usefulness for faculty calibration or providing objective feedback in a virtual setting.

TABLE III: Descriptive Statistics for Faculty Survey

| Response variable	Chi-square	*p*-value	Mean group TA	Mean group 3D	Mean group 3D+	Pairwise p < 0.05
Ease of use	6.6252	*0.03642*	1.571	2.166	3.333	3D+ vs. TA
Time efficiency	6.3774	*0.04123*	1.857	1.666	3.500	3D+ vs. 3D (marginally sig)
Provides constructive feedback for students	1.7115	0.425	1.571	1.666	1.166	NS*
Useful for faculty calibration	1.8872	0.3892	1.571	1.500	1.166	NS
Provides objective feedback in virtual setting	4.5418	0.1032	2.571	1.666	1.500	NS

Note: *Not significant.
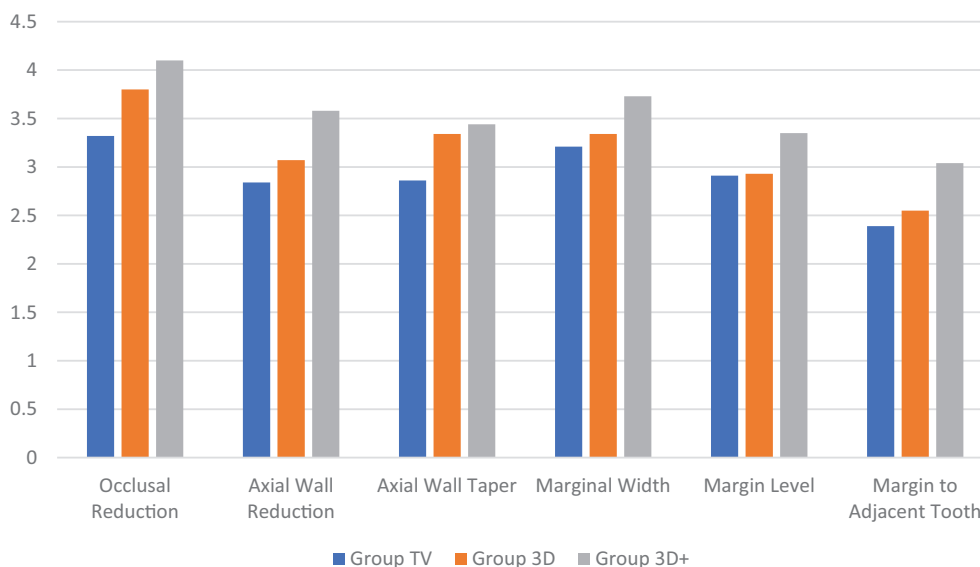
Fig. 2. Means grades per group for criteria evaluated.

TABLE IV: Descriptive Statistics for Faculty Grades

Response variable	LR p-value	Group TA	Group 3D	Group 3D+
Occlusal reduction	0.06231	3.32 ± 0.62	3.80 ± 0.70	4.10 ± 0.38
Axial wall reduction	*0.04279*	2.84 ± 0.36	3.07 ± 0.75	3.58 ± 0.41
Axial wall taper	0.2184	2.86 ± 0.49	3.34 ± 0.83	3.44 ± 0.57
Marginal width	0.05152	3.21 ± 0.33	3.34 ± 0.34	3.73 ± 0.51
Margin level	0.5317	2.91 ± 0.60	2.93 ± 0.82	3.35 ± 0.95
Margin to adjacent tooth	0.3791	2.39 ± 0.68	2.55 ± 0.58	3.04 ± 1.26

Fig. 2 depicts mean grades per group for each evaluated criterion. The 3D+ scores show a higher trend than groups 3D and TV. Table IV shows the descriptive statistics, including LR p-value, mean value, and standard deviation for Preparation Grading. Results indicated that only axial wall reduction differed significantly from the three assessment methods. 3D+ faculty graders scored axial wall reduction significantly higher than the TV faculty ($R2 = 0.2704703$). Occlusal reduction ($R2 = 0.2461607$) and marginal width ($R2 = 0.2586628$) showed marginally significant trends, where the 3D+ faculty scored the variables higher than the TV faculty.

4. Discussion

This pilot study explored the effectiveness of using computerized evaluation tools for grading student simulation exercises in preclinical dental education. The results of this pilot study confirmed the null hypothesis with respect to most grading parameters except axial wall reduction. There was no significant difference between 3D Group and TV Group scores, which coincides with a previous study by Lee *et al.* [12]. The null hypothesis for axial wall reduction was rejected, as it showed significantly higher scores for the 3D+ Group compared to the TV Group. The authors suggest this may be due to the Compare® software's ability to measure differences between the MP and the student preparation. It is important to mention that for both TV and 3D groups, no matrix was used to gauge the axial

reduction. The Compare® software was able to measure from the finish line toward the occlusal surface. Other measurements, such as distances between the preparations and the adjacent teeth and the gingival crest were not measurable by the software. These limitations were confirmed by Callan *et al*. [15]. Sly *et al*. eliminated certain parameters in grading due to limitations of the Compare software as well; however, this current study addressed the software limitations through the addition of a 3D scanned model assessment [9].

Although faculty survey results showed no significant differences among the three assessment methods with respect to usefulness in providing constructive feedback to students, usefulness for faculty calibration, and providing objective feedback in a virtual environment, there was a positive trend in the responses. On the other hand, the ease of use and efficiency of the three assessment methods varied significantly. Faculty did not have to scan the preparations to reduce variables in this pilot study. Still, they were calibrated using the software for visual assessment and for comparing the preparations to the MP. Despite this, faculty still required additional support. Utilization of Compare® software was more difficult and less efficient than traditional grading, which may pose a barrier to its use as a standard assessment tool. With a steep learning curve, the more faculty became comfortable with digital software, the more efficient they became, as seen in a study by *Shih et al*. [16]. Until then, both initial learning of this technology and its application in grading student projects continues to be a challenge [9].

Limitations of this pilot study included a small sample size of preparations. The investigators selected these samples as they embodied a wide variety of common errors noted among novice operators. Another limitation, not specific to this pilot study but in general use of technology as assessment tools, is the lack of in-person student-faculty interactions. These results provide a foundation for further exploration of these computerized evaluation tools in preclinical dental education. The investigators suggest that these digital tools enhance objectivity and generate conversations between the evaluator and the learner. For example, when assessing student projects remotely, the evaluator can schedule a virtual meeting with the learner. This enables the learner to visualize their preparations as a three-dimensional scan and discuss improvement plans. This process emphasizes an interactive teaching-learning opportunity based on accurate and objective data, magnified, rotated and manipulated on the screen to highlight fundamental restorative and prosthodontic concepts. Positive attributes and areas for improvement can be pointed out for reflective learning. Additionally, the interpersonal aspect of teaching and learning is not removed. This pilot study suggests that the integration of 3D model visual assessment and comparative analysis software reports can provide a more comprehensive evaluation of student full-coverage crown preparations than either method alone. Further research is needed to expand the use of these tools in dental education and to address the limitations of this pilot study.

5. Conclusion

This pilot study concludes that comparative analysis software-generated reports and 3D model visual grading combined can overcome the limitations of traditional evaluation methods and digital methods alone by providing objective feedback to students and enhancing their learning experience both in person and remotely.

Conflict of Interest

The authors declare that they do not have any conflict of interest.

References

[1] Abdalla R, Bishop SS, Villasante-Tezanos AG, Bertoli E. Comparison between students' self-assessment, and visual and digital assessment techniques in dental anatomy wax-up grading. *Eur J Dent Educ*. 2021;25(3):524–35. doi: 10.1111/eje.12628.

[2] Taylor CL, Grey NJA, Satterthwaite JD. A comparison of grades awarded by peer assessment, faculty and a digital scanning device in a pre-clinical operative skills course. *Eur J Dent Educ*. 2013;17(1):e16–21. doi: 10.1111/j.1600-0579.2012.00752.x.

[3] Patel SA, Barros JA, Clark CM, Frey GN, Streckfus CF, Quock RL. Impact of technique-specific operative videos on first-year dental students' performance of restorative procedures. *J Dent Educ*. 2015;79(9):1101–7. doi: 10.1002/j.0022-0337.2015.79.9.tb06004.x.

[4] Schepke U, van Wulfften Palthe ME, Meisberger EW, Kerdijk W, Cune MS, Blok B. Digital assessment of a retentive full crown preparation—An evaluation of prepCheck in an undergraduate pre-clinical teaching environment. *Eur J Dent Educ*. 2020;24(3):407–24. doi: 10.1111/eje.12516.

[5] Chiang H, Staffen A, Abdulmajeed A, Janus A, Bencharit CS. Effectiveness of CAD/CAM technology: a self-assessment tool for preclinical waxing exercise. *Eur J Dent Educ*. 2021;25(1):50–5. doi: 10.1111/eje.12576.

[6] Park CF, Sheinbaum JM, Tamada Y, Chandiramani A, Lian L, Lee C, *et al.* Dental students' perceptions of digital assessment software for preclinical tooth preparation exercises. *J Dent Educ*. 2017;81(5):597–603. doi: 10.21815/jde.016.015.

[7] Miyazono S, Shinozaki Y, Sato H, Isshi K, Yamashita J. Use of digital technology to improve objective and reliable assessment in dental student simulation laboratories. *J Dent Educ*. 2019;83(10):1224–32. doi: 10.21815/jde.019.114.

[8] Callan RS, Palladino CL, Furness AR, Bundy EL, Ange BL. Effectiveness and feasibility of utilizing E4D technology as a teaching tool in a preclinical dental education environment. *J Dent Educ*. 2014;78(10):1416–23. doi: 10.1002/j.0022-0337.2014.78.10.tb05815.x.

[9] Sly MM, Barros JA, Streckfus CF, Arriaga DM, Patel SA. Grading class I preparations in preclinical dental education: e4D compare software vs. the traditional standard. *J Dent Educ*. 2017;81(12):1457–62. doi: 10.21815/jde.017.107.

[10] Renne WG, McGill ST, Mennito AS, Wolf BJ, Marlow NM, Shaftman S, *et al.* E4D compare software: an alternative to faculty grading in dental education. *J Dent Educ*. 2013;77(2):168–75. doi: 10.1002/j.0022-0337.2013.77.2.tb05459.x.

[11] Hamil LM, Mennito AS, Renné WG, Vuthiganon J. Dental students' opinions of preparation assessment with E4D compare software versus traditional methods. *J Dent Educ*. 2014;78(10):1424–31. doi: 10.1002/j.0022-0337.2014.78.10.tb05816.x.

[12] Lee C, Kobayashi H, Lee SR, Ohyama H. The role of digital 3D scanned models in dental students' self-assessments in pre-clinical operative dentistry. *J Dent Educ*. 2018;82(4):399–405. doi: 10.21815/jde.018.046.

[13] Siddanna GD, Karpenko AE, Mantesso A, Sterlitz SJ, Fasbinder DJ. Virtual clinic: applying technology to expand dental school walls. *J Dent Educ*. 2022 Mar 1; 408–14. doi: 10.1002/jdd.13119.

[14] R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *Published Online*. 2020. Accessed May 21, 2023. Available from: https://www.R-project.org/.

[15] Callan RS, Van Haywood B, Cooper JR, Furness AR, Looney SW. The validity of using E4D compare's % comparison to assess crown preparations in preclinical dental education. *J Dent Educ*. 2015;79(12):1445–51. doi: 10.1002/j.0022-0337.2015.79.12.tb06044.x.

[16] Shih W, Tran K, Yang V, El Masoud B, Sexton C, Zafar S. Investigation of inter-and intra-rater reliability using digital dental software for prosthodontics crown preparations. *J Dent Educ*. 2020;84(9):1037–45. doi: 10.1002/jdd.12180.