Forum for Education Studies 2024, 2(4), 1609.
https://doi.org/10.59400/fes1609

Article

Intra and inter-individual differences in the visual evaluation of crown preparations in the phantom head

Maximilian Nothaft[1,2,*], Laurenz Kotthaus[1], Eva Groth[1], Mihai Rominu[3], Rüdiger Junker[1]

[1] Department of Prosthodontics and Biomaterials, Faculty of Medicine and Dentistry, Danube Private University, Steiner Landstrasse, 3500 Krems-Stein, Austria
[2] Dental Practice, Dres Marion & Maximilian Nothaft, 94032 Passau, Germany
[3] Department of Protheses Technology and Dental Materials, "Victor Babes" University of Medicine and Pharmacy, 300041 Timisoara, Romania
*** Corresponding author:** Maximilian Nothaft, Maximilian.nothaft@dp-uni.ac.at

CITATION

Nothaft M, Kotthaus L, Groth E, et al. Intra and inter-individual differences in the visual evaluation of crown preparations in the phantom head. Forum for Education Studies. 2024; 2(4): 1609. https://doi.org/10.59400/fes1609

ARTICLE INFO

Received: 12 August 2024
Accepted: 14 October 2024
Available online: 19 November 2024

Abstract: Objectives: This study aimed to evaluate the intra- and inter-individual differences between the first and second visual assessments of crown preparations made in the phantom head by pre-clinical undergraduates. **Material and Methods:** Third-semester dental students at the Danube Private University conducted a preparation on a model tooth to accommodate a crown made of IPS e. max CAD material. At regular intervals of three weeks, 20 crown preparations were evaluated on two occasions. The crown preparations were evaluated by 3 students using a visual approach and predetermined parameters, following a "noise audit. The intra- and inter-rater reliabilities of the method were evaluated using Fleiss and Cohen's kappa. **Results:** Results showed an average fair variability between the assessors in the first ($\kappa = 0.29$) and second ($\kappa = 0.22$) visual individual assessment using the predetermined parameters and a performed noise audit. Rater 1 ($\kappa = 0.47$) and Rater 2 ($\kappa = 0.49$) showed moderate variability and Rater 3 ($\kappa = 0.33$) fair variability in the mean intra-individual variability by individual rating. **Conclusion:** Within the limitations of this study, intra- and inter-individual variabilities were observed, although the assessment criteria and prior 'noise audits' were defined. Further studies with bigger sample size and longer durations are required to identify methods to reduce this variability.

Keywords: prep Check; dental education; self-assessment; pre-clinical preparation course

1. Introduction

One of the primary objectives of dental education is to provide pre-clinical students with practical training. To achieve this, it is essential to evaluate the preparations conducted on the phantom head, which is a training dummy for dental students, where practical exercising takes place. This is typically performed by assessors, who are primarily senior clinicians, preferably consultants. Such senior clinicians and consultants are highly trained and therefore not available in high quantities and needed in patient treatment. We decided to use student tutors in higher semesters which passed this course already to support the senior clinicians during practical training.

However, ensuring the reliability of such assessments is challenging [1,2]. The assessment encompasses two distinct types of variability. The first is intra-individual variability, which is broadly defined as fluctuations in an individual's cognitive performance over time [3]. The second factor to be considered is inter-individual variability, which refers to the discrepancy in assessments between two or more assessors [4]. It is of paramount importance that the variability between different

assessors and between one examiner at different times is kept to an absolute minimum to guarantee reliable evaluation [5]. Although the importance of this topic has been acknowledged, available literature is limited.

Several approaches can be used to overcome these variations. The literature shows various attempts on this topic. Here the simplest method is the evaluation following the "glance and grade" method [6]. To make results more comparable a standardized process is recommended. Here clearly stated criteria and checklists should be used. Also, a regular assessment training of the raters can reduce assessment variance. Kahneman et al. stated the use of "noise audits" for this purpose. The term "noise" the authors describes the blurring of ratings between different raters and by the same rater at different times [1,7].

With the integration of intraoral scanners, a further possibility of standardization is given. With special analyzing software such as prep Check (Dentsply Sirona Global Headquarters, Charlotte, USA) preparations can be scanned and analyzed with these tools without human interference such as Interrater Variability. It also gives assessors the possibility to train their vision for correct tooth preparations whereas students get the possibility to evaluate their work without a clinician since it is easy to use and gives clear visual presentation of the criteria using color coding [7–16].

Considering this background, the authors investigated the reliability of such examinations using a newly designed study setup using such "noise audits" in combination with clear assessment guidelines to reduce variance among raters.

Against the background of personnel shortage especially by consultants the consideration was that trained students in higher semesters (≥ 10th semester) who already passed the preparation course and already gained some clinical experience during their patient treatment could partly take over the task of assessing crowns during the preparation course to support senior clinicians and consultants.

This study aimed to investigate the intra- and inter-individual differences between the first and second visual evaluations of student preparations of a molar in the Frasaco model for full coverage with an all-ceramic crown carried out by trained students in the 10th semester of dental school. It was tested whether the students can achieve reproducible assessment and whether the assessment variability can be reduced by calibrating the assessors repeatedly.

It was hypothesized that despite a prior determination of guidelines in a "noise audit," intra- and inter-individual variability occurs in assessment. It was also hypothesized that students are not well suited to assess preparations and can there for not be used as assessors.

2. Material and methods

The third dental semester students at Danube Private University (DPU) were instructed to produce a fully anatomical monolithically milled crown using Cerec (Dentsply Sirona Global Headquarters, Charlotte, USA).

Each student was given the option to provide their prepared teeth to the assessors and participate in this project. This experimental study comprised 20 prepared teeth obtained from 20 students (n = 20). The Ethics Committee of Danube Private

University had no concerns about conducting this study (Ethical voting number: GZ: DPU-EK/062).

The third semester students which prepared the teeth and the 10th semester student assessors received pre-defined criteria for evaluating the prepared tooth stumps, ensuring an assessment. (**Figure 1**, **Table 1**) These criteria were developed by the head of the Department of Prosthodontics and Biomaterials at the Danube Private University, Krems a. t. D., Austria following actual recommendations in the literature [17].

The assessors followed these guidelines combined with manufacturer instructions for assessing the preparation of IPS e. max CAD single-tooth crowns, as shown in **Figure 1** and **Table 1**.

The student assessors received an assessment training before this evaluation by a senior clinician. The decision on whether a criterion could be considered passed was made based on a numerical evaluation. If the required conditions were not met, the task was considered a failure. The task was immediately deemed to have failed if a neighboring tooth was damaged.

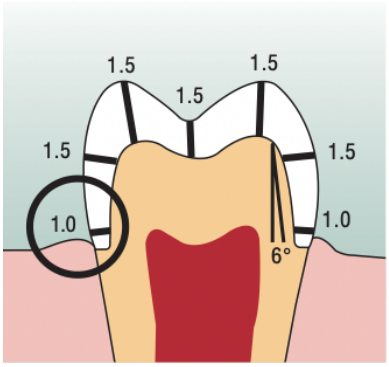

Figure 1. Preparation guidelines for e. max CAD FDP.
Source: ivoclar vivadent AG.

Table 1. pre-defined criteria for evaluation of prepared tooth stumps the table shows the values that must be achieved to fulfill the criterion.

Criterion	Value to be fulfilled	Tolerance
preparation angle	3–6°	+3°
undercuts	No undercut areas	More than 60% of the surface matches the 'fulfilled' criteria
occlusal reduction	1.5 mm	+0.5 mm
distance to antagonist	1.5 mm	+0.5 mm
axial reduction	1.5 mm	+0.5 mm
surface roughness	Smooth surface, no sharp edges	More than 60% of the surface matches the 'fulfilled' criteria
preparation margin	1.0 mm	+0.2 mm
marginal finish type	Chamfer	More than 60% of the surface matches the 'fulfilled' criteria
clinical usability evaluation	Preparation is clinically usable	Little corrections are necessary to fulfill the abovementioned criteria

The tolerance values up to which a criterion can still be assessed as `fulfilled´ can be found opposite.

The assessors conducted initial evaluations individually and objectively, adhering to the preparation parameters and corresponding tolerance limits. Beforehand, they conducted a 'noise audit' where the assessors got instructed by a senior clinician in the rules of preparation assessment and hereby calibrated. They also discussed the evaluation criteria of student preparations and defined tolerance ranges among each other's. The evaluation of five example preparations was also practiced under supervision. All three assessors were students in the clinical part of the dental education [7]. This 'noise audit' provides instructions to the three assessors on the correct application of the evaluation criteria to ensure fairness and the practicing should minimize inter-individual evaluation discrepancies. To evaluate the prepared dies, a millimeter-scale probe and a silicone ridge (Silaplast, Detax GmbH, Ettlingen, Germany) that had been previously molded over an unprepared 36 Frasaco tooth were used. (**Figure 2**) Each parameter was tested and assessed as either 'fulfilled' (within the tolerance range) or 'not fulfilled' (outside the tolerance range).

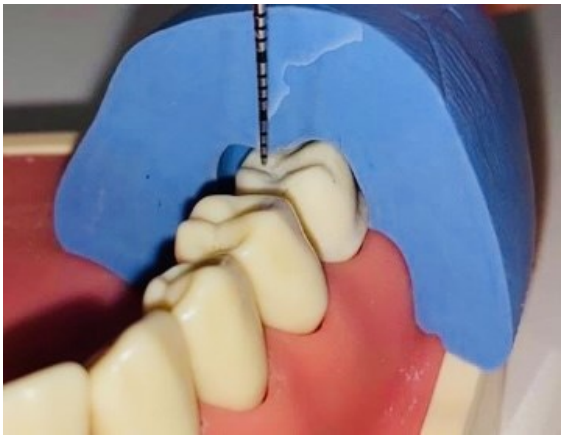

Figure 2. Silicone pre-cast with millimeter-scale measuring probe here the occlusal reduction of a model tooth prepared following the guidelines for IPS e. max CAD single tooth crowns is measured.

A second visual assessment was conducted three weeks after the initial assessment under the same conditions. The purpose the three-week interval was chosen was to prevent any influence of the previous assessment on the second assessment owing to the time interval. First, an intra-individual comparison was performed by comparing the evaluations of the same assessor in the first and second evaluation sessions.

After respectively the first and second visual assessment a group assessment was additionally carried out. Here the assessors discussed why they marked a criterion of each preparation as whether fulfilled or not. This peer evaluation was done so that the assessors could compare their evaluation of the preparations to the others and a possibility to discuss their point of view was given. This should help the assessors to achieve better and homogenous results.

Second, an inter-individual comparison was conducted by comparing the first and second evaluations by all three assessors.

The collected data were statistically analyzed using Cohen's and Fleiss´ kappa values.

2.1. Cohen's kappa

As previously stated, the Cohen's kappa calculation can be used to compare intra-individual variability, specifically, the agreement of a tester on different test days.

The result of the equation is indicated by kappa (= κ), with collected results ranging between −1 and +1, indicating the level of unanimity. A result of ≤ 0 indicates poor agreement, while a result between 0.0 and 0.2 is considered low agreement. Results between 0.21 and 0.4 indicate fair agreement. According to Landis and Koch [13], a result of 0.41–0.6 indicates moderate agreement, while a result of 0.61–0.8 indicates substantial agreement. A result between 0.81–1 indicates almost perfect agreement.

The calculation employed a dual system that recognizes only two digits (0 and 1). The point 'fulfilled' is expressed as 1 and 'not fulfilled' as 0. The overall score was calculated similarly. A score of 1 indicated that the student passed, whereas a score of 0 indicated that the student failed [13].

2.2. Fleiss kappa

This calculation method can be used to determine the level of agreement between assessors for specific assessment parameters. This is particularly useful when more than two inspectors are involved in an evaluation.

Fleiss Kappa is a measure of agreement that ranges from −1 to +1. Classification of the level of agreement was identical to that of Cohen's kappa, as described above [13].

3. Results

The assessors' individual visual assessments were compared to determine intra- and inter-individual differences.

3.1. Intra-individual agreement of the assessors

Regarding the first evaluation criterion, 'preparation angle', the comparison of the first and second evaluation rounds for the different assessors resulted in Cohen's kappa values (κ) of $\kappa = 0.29$ for assessor 1, indicating moderate agreement of the evaluation results between the first and second evaluation rounds. Assessor 2 had a value of $\kappa = 0.6$, also indicating moderate agreement. Assessor 3 demonstrated poor agreement between the results of the two assessment rounds, with a negative value of $\kappa = -0.01$.

The second evaluation criterion 'undercuts' showed a κ value of 0.79 for assessor 1 and a value of 0.68 for assessor 2, indicating a substantial agreement of the evaluation results between rounds 1 and 2. Grader 3 showed only fair agreement with a value of 0.21. For the 'occlusal reduction' assessment, assessor 1 achieved moderate agreement with $\kappa = 0.6$, assessor 2 moderate agreement with $\kappa = 0.23$ and assessor 3 poor agreement with $\kappa = -0.07$ between the results of the two assessment rounds.

Testers 2 and 3 achieved a perfect agreement value of $\kappa = 1$ for the parameter 'distance to antagonist' in evaluation rounds 1 and 2. However, tester 1 had a κ value of −0.05. Tester 1 achieved a moderate agreement value of 0.53 for 'axial removal', while tester 2 and tester 3 achieved κ values of 0.04 and 0.2 respectively, indicating low agreement. The 'surface roughness' was also evaluated by multiple testers on different days. Tester 1 had substantial agreement between the first and second evaluation day, while testers 2 and 3 had moderate agreement ($\kappa = 0.53$ and $\kappa = 0.49$).

Assessors assessed the 'preparation margin' with varying levels of agreement: $\kappa = 0.4$ for assessor 1, $\kappa = 0.5$ for assessor 2, and $\kappa = 0.38$ for assessor 3.

Similarly, the 'margin finish type' was assessed with Cohen's kappa values of 0.5 for assessor 1, 0.35 for assessor 2, and 0.44 for assessor 3. (**Table 2**)

The results of the visual evaluation were then summarized with respect to the 'fulfilled' and 'not fulfilled' evaluation parameters. If over 60% of the criteria were met, the student's preparations were evaluated as either 'passed' or 'failed'.

Table 2. Overview of the Cohen´s Kappa values of the various assessment parameters of the three assessors.

Criteria	κ-value of intra-individual variability of assessor 1	κ-value of intra-individual variability of assessor 2	κ-value of intra-individual variability of assessor 3
preparation angle	0.29	0.6	−0.01
undercuts	0.79	0.68	0.21
occlusal reduction	0.6	0.23	−0.07
distance to antagonist	−0.05	1	1
axial reduction	0.53	0.04	0.2
surface roughness	0.68	0.53	0.49
preparation margin	0.4	0.5	0.38
marginal finish type	0.5	0.35	0.44

3.2. Inter-individual agreement of the assessors

After analyzing the difference among the testers intra-individual assessment, the inter-individual agreement was tested using Fleiss-Kappa as a statistical tool.

For the first criteria 'preparation angle' there was moderate agreement between the three assessors in the first assessment round, with a value of $\kappa = 0.25$, but poor agreement in the second round, with $\kappa = -0.02$. The first visual assessment of the criterion 'undercuts' showed moderate inter-rater agreement ($\kappa = 0.46$). The second visual assessment showed a fair agreement ($\kappa = 0.36$). The assessors reached more similar results in the first than in the second visual assessment. Looking on the results for 'occlusal reduction', the first individual visual assessment showed low inter-rater agreement ($\kappa = 0.06$). This was followed by a low level of agreement ($\kappa = 0.16$). First and second individual visual evaluations for 'distance to antagonist' showed poor agreement among the testers, with κ values of −0.03 and −0.02, respectively. The first

individual visual evaluation of the criterion 'axial reduction' showed low agreement between assessors ($\kappa = 0.10$), and the second evaluation showed a similar result ($\kappa = 0.17$). Concerning the 'surface roughness', the first individual visual evaluation showed substantial agreement ($\kappa = 0.69$), while the second evaluation, which took place three weeks later, showed moderate agreement ($\kappa = 0.38$). The agreement between the assessors was higher in the first evaluation than in the second. The assessors showed moderate agreement ($\kappa = 0.46$) in their first evaluation of the preparation margin when assessing the 'preparation margin'. This was repeated in the second evaluation ($\kappa = 0.41$).

Regarding the evaluation criterion 'margin finish type', the first individual visual evaluation resulted in fair inter-rater agreement ($\kappa = 0.36$). The second individual visual evaluation, conducted three weeks after the first one, also resulted in fair agreement ($\kappa = 0.30$), as did the first evaluation. (**Table 3**).

Table 3. Overview of the Fleiß-Kappa values of the various assessment parameters in the 1st and 2nd visual assessment.

Criteria	κ-value of inter-individual variability of the first visual evaluation	κ-value of inter-individual variability of the second visual evaluation
preparation angle	0.25	−0.02
undercuts	0.46	0.36
occlusal reduction	0.06	0.16
distance to antagonist	0.03	0.02
axial reduction	0.10	0.17
surface roughness	0.69	0.38
preparation margin	0.46	0.41
marginal finish type	0.36	0.30

4. Discussion

In this study, we investigated the intra- and inter-individual variability in the visual evaluation of crown preparations in the phantom head. It was hypothesized that despite a prior determination of guidelines in a "noise audit" intra- and inter-individual variability occurs in assessment.

The assessment variability between individual assessors at different assessment times and between different assessors was clear for the parameters in the first and second individual assessments, although attempts were made to reduce these differences using pre-defined criteria and a prior noise audit, as described by Kahneman et al. [7].

These findings corroborated the initial hypothesis. Within the limitations of this study, it was demonstrated that the pre-defined criteria and noise audits could not standardize the assessments.

There are different reasons why assessment variability occurs. Kahneman describes that individuals are controlled by emotions, which can bias their judgments. There is also a bias due to subjective weighting of criteria between different assessors

or individual preferences. He also states that people often overestimate their ability to make precise predictions where experienced assessors in particular tend to trust their intuition too much [15,18]. Therefore, assessors' judgments cannot be reproduced visually and are subject to variation.

Also, fatigue is discussed in the work of Al Amri et al. as possibility for inequality in assessment. Here the authors mention that there is no significant different in the results of the assessment [19]. Other research brings up that the gain of experience while performing a task can lead to different assessment results.

Cognition refers to processes, such as perception, attention, thinking, and memory. Emotions determine whether we can remember experienced events or judgments, and the decisions we make [18,20].

This reflects the findings of this paper.

A dissertation by Baumann [10] showed similar results. The assessors achieved a lower level of agreement when assessing by visual inspection without fixed criteria just according to their educational/scientific background than when assessing using fixed assessment criteria [10]. This finding is also confirmed by Al Amri et al. [19].

Alammari et al. [1] reported similar results when investigating the variability in impression procedures carried out by undergraduates. The authors found significant inconsistencies among the assessors when evaluating using the glance and grade methods. Similar findings were reported by Lilley et al., Fuller, Salvendy et al., and Jenkins et al. [6,15,20,21].

To reduce variability, Goepferd and Kerber [22] used an analytical system with specific evaluation criteria similar to those used in this study. They observed a reduction in variability, which contradicts our findings, but is like the findings of Schmitt et al. [22,23]. This study also used predefined criteria to evaluate tooth preparation. The assessors were given clear criteria which they used to evaluate all preparations. This was made to reduce the variability between different raters but also between the same rater at different points in time. To reduce variability even more a 'noise audit' was integrated before evaluation to clarify the assessment parameters and train the assessors. Although the assessment was carried out with such an analytical grading system and a 'noise audit' high difference in the evaluation was found.

To reduce the level of variability among assessors, the use of digital technology, such as prep Check, has also been discussed. Prep Check allows the assessor to measure every point of a tooth preparation and compare it either to a Master preparation or against predefined parameter. It is independent in its decision making and visualizes the result of the assessment clearly by using a color scale. This makes it easy for the assessor to make his decision regarding grading the preparation [24–27]. Whereas students state that the feedback of a senior clinician is necessary to improve individual performance and cannot completely be replaced with digital solution although they seem to be less reliable. But for a quick and unbiased assessment such software as prep Check is ideal and supports both students and assessors. But when used during examination a higher rate of students do not pass the assessment due to stricter judgment of prep Check compared to human assessors [28–30].

Although the impact of our study is limited due to the small number of assessed teeth, the low experience level of the raters and a potential bias because of the study set up which did not take place in exam like circumstances we could clearly show that

the assessment of preparations is not as reliable as we might think. This gives us the chance to investigate this topic further also because especially new literature from the past five years to this topic is limited. The influence of the new aspect of assessment criteria clarification and assessment training with the established 'Noise Audit' might be a very interesting point to investigate. The study was set up with students as assessors. In further investigations we are planning to use senior clinicians as assessors and then compare the findings to see whether students are as good as senior clinicians in means of preparation assessment.

5. Conclusion and future recommendations

Despite the established grading guidelines and a prior 'noise audit' between the assessors, high intra- and inter-individual differences were observed in the visual grading of the test subjects.

We hereby can conclude that student tutors in high semesters are not well suited to assess crown preparations reproducible. Even though we tried to reduce the variability by calibrating the assessors we were not able to reduce the variability to a sufficient extent. So, they cannot support the senior clinicians and consultants in preparation assessment and should only be used for supporting the lower grade students by giving instructions to fulfil their given tasks.

Further investigation is needed on this topic to identify the reasons for assessor variability and methods to overcome it. We suggest studies with a bigger sample size and longer duration to get appropriate findings. We are also conducting a study where we compare the inter and intraindividual variability of senior clinicians with a bigger sample size. It is important to create reliable assessment systems for student education. Before this background the use of technical solution such as prep Check seems to be adequate. Such tools can help the assessors when used in assessment training to reduce variability. It can be investigated in further studies whether the use of prep Check during examination is practicable.

Author contributions: Conceptualization, RJ and MN; methodology, RJ, LK and MN; software, MN and LK; investigation, MN and LK; data curation, MN and LK; writing—original draft preparation, MN and LK; writing—review and editing, MR and EG; visualization, MN; supervision, RJ and MR. All authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

References

1. Alammari, M.R., Y.M. Alkhiary, and A.A. Nawar, Intra-and inter-examiner variability in evaluating impression procedures at the undergraduate level. Journal of Life Sciences, 2013. 5(1): p. 5-10.
2. Sharaf, A.A., A.M. AbdelAziz, and O.A. El Meligy, Intra- and inter-examiner variability in evaluating preclinical pediatric dentistry operative procedures. J Dent Educ, 2007. 71(4): p. 540-4.
3. Haynes, B.I., S. Bauermeister, and D. Bunce, Age and Intraindividual Variability, in Encyclopedia of Geropsychology, N.A. Pachana, Editor. 2015, Springer Singapore: Singapore. p. 1-9.
4. Faure, P., et al., Social Determinants of Inter-Individual Variability and Vulnerability: The Role of Dopamine. Front Behav Neurosci, 2022. 16: p. 836343.

5. Miyazono, S., et al., Use of Digital Technology to Improve Objective and Reliable Assessment in Dental Student Simulation Laboratories. J Dent Educ, 2019. 83(10): p. 1224-1232.
6. Jenkins, S.M., et al., Evaluating undergraduate preclinical operative skill; use of a glance and grade marking system. J Dent, 1998. 26(8): p. 679-84.
7. Kahneman, D., O. Sibony, and C.R. Sunstein, Noise : a flaw in human judgment. First edition. ed. 2021, New York: Little, Brown Spark. ix, 454 pages.
8. Kwon, S.R., et al., Dental anatomy grading: comparison between conventional visual and a novel digital assessment technique. J Dent Educ, 2014. 78(12): p. 1655-62.
9. Kateeb, E.T., et al., Utilising an innovative digital software to grade pre-clinical crown preparation exercise. Eur J Dent Educ, 2017. 21(4): p. 220-227.
10. Baumann, M., Evaluation von Bewertungskriterien für praktische Studentenarbeiten im Vergleich zur Bewertung per Augenschein. 2015, LMU München.
11. Schiefelbein, R., Untersuchung zur Umsetzung von Richtlinien zur Präparation CAD/CAM-generierter vollkeramischer Frontzahnkronen. 2015, LMU München.
12. Nothaft, M., et al., The preclinical teaching of the "Chairside Economical Restoration of Esthetic Ceramic" workflow: A questionnaire-based evaluation. Saudi Journal of Oral Sciences, 2023. 10(2): p. 72-77.
13. Landis, J.R. and G.G. Koch, The measurement of observer agreement for categorical data. Biometrics, 1977. 33(1): p. 159-74.
14. Welk, A., et al., Computer-assisted learning and simulation systems in dentistry--a challenge to society. Int J Comput Dent 2006. 9: p. 253-265.
15. Salvendy, G., et al., Pilot study on criteria in cavity preparation--facts or artifacts? J Dent Educ, 1973. 37(11): p. 27-31.
16. Mino, T., et al., Rating criteria to evaluate student performance in digital wax-up training using multi-purpose software. J Adv Prosthodont, 2022. 14(4): p. 203-211.
17. Strub, J.R., et al., Curriculum Prothetik: Band 1. 2019: Quintessenz Verlag.
18. Brosch, T., et al., The impact of emotion on perception, attention, memory, and decision-making. Swiss Med Wkly, 2013. 143: p. w13786.
19. Al Amri, M.D., H.R. Sherfudhin, and S.R. Habib, Effects of Evaluator's Fatigue and Level of Expertise on the Global and Analytical Evaluation of Preclinical Tooth Preparation. J Prosthodont, 2018. 27(7): p. 636-643.
20. Fuller, J.L., The effects of training and criterion models on interjudge reliability. J Dent Educ, 1972. 36(4): p. 19-22.
21. Lilley, J.D., et al., Reliability of practical tests in operative dentistry. Br Dent J, 1968. 125(5): p. 194-7.
22. Goepferd, S.J. and P.E. Kerber, A comparison of two methods for evaluating primary class II cavity preparations. J Dent Educ, 1980. 44(9): p. 537-42.
23. Schmitt, L., et al., Study on the Interrater Reliability of an OSPE (Objective Structured Practical Examination) - Subject to the Evaluation Mode in the Phantom Course of Operative Dentistry. GMS J Med Educ, 2016. 33(4): p. Doc61.
24. Habib, S.R. and H. Sherfudhin, Students' self-assessment: a learning tool and its comparison with the faculty assessments. J Contemp Dent Pract, 2015. 16(1): p. 48-53.
25. Schepke, U., et al., 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): p. 407-424.
26. Wolgin, M., et al., Comparison of a prepCheck-supported self-assessment concept with conventional faculty supervision in a pre-clinical simulation environment. Eur J Dent Educ, 2018. 22(3): p. e522-e529.
27. Stoilov, M., et al., Comparison of Digital Self-Assessment Systems and Faculty Feedback for Tooth Preparation in a Preclinical Simulation. Int J Environ Res Public Health, 2021. 18(24).
28. Nagy, Z.A., et al., Evaluating the efficiency of the Dental Teacher system as a digital preclinical teaching tool. Eur J Dent Educ, 2018. 22(3): p. e619-e623.
29. Nothaft, M., et al., Self-Assessment-Of-Dental-Preparations-In-The-Phantom-Head-With-Prepcheck-A-Questionnaire-Based-Evaluation. Romanian Journal of Oral Rehabiltation, 2023.
30. Schlenz, M.A., et al., Undergraduate dental students' perspective on the implementation of digital dentistry in the preclinical curriculum: a questionnaire survey. BMC Oral Health, 2020. 20(1): p. 78.