Establishing Evaluation Criteria for Digital Dental Simulation Systems and the Differences between Software and Evaluators

Tokunaga J[1], Morishita M[1]*, Muraoka K[1], Kibune R[1], Yoshii S[2] and Awano S[1]

[1]Division of Clinical Education Development and Research, Department of Oral Function, Kyushu Dental University, Japan

[2]Division of Promoting Learning Design Education, Department of Physical Function, Kyushu Dental University, Japan

*Corresponding author: Masaki Morishita, DDS, PhD, Division of Clinical Education Development and Research, Department of Oral Function, Kyushu Dental University, 2-6-1 Manazuru, Kokurakita, Kitakyushu 803-8580, Japan, Tel: +81-93-582-1131; Fax: +81-93-582-1140; Email: r08morishita@fa.kyu-dent.ac.jp

Abstract

Purpose: Cavity and tooth preparation are essential curriculum components in dental education for undergraduate students, and faculty staff determine whether the standard of the work is acceptable. However, evaluators are inevitably subject to variability and bias, so quantification of the evaluation is desirable. Therefore, various digital devices have been developed to quantify evaluations, such as SURFLACER, Simodont, and DentSim. In this study, we aimed to investigate the use of evaluation software and to compare human evaluation with software evaluation using a new digital dental education device with a real-time evaluation function.

Material and Method: Dental students prepared a full metal crown on an artificial mandibular left first molar, which was evaluated by the software and a human evaluator (a dentist).

Results: The results showed that the software evaluation was more suitable with eight cross-sections than with four cross-sections. The dentist evaluated the same teeth according to a set of evaluation criteria, but there were no statistically significant differences in any of the evaluation items. Finally, multivariate analysis was conducted to assess the relationship between the software evaluation using eight cross-sections and the human evaluator using dummy variables, and it revealed that the margin form evaluation by the dentist was a significant explanatory variable for the score by the software.

Conclusion: This preliminary study is the first to demonstrate that the new digital dental education device has considerable potential as a digital educational tool.

Keywords: Dental Digital Simulation System; Real-time evaluation; Tooth Preparation

Abbreviations: 3D: Three-dimensional; CT: Computed Tomography; CAD: Computer Aided Design; CAM: Computer Aided Manufacturing; SD: Standard Deviation.

Introduction

Many individuals are now in possession of digital devices such as smartphones and can easily handle digital data. It has become possible for individuals to create three-dimensional (3D) models using 3D printers. In the field of dentistry, 3D models of the jawbone using computed tomography (CT) imaging data can be used for diagnosis and surgery, and same-day fabrication of computer aided design / computer aided manufacturing (CAD/CAM) crowns using optical impressions and milling equipment is being used by general practitioners [1]. Digitalization of the dental education field is also becoming more imperative to keep up with current trends. The Model Core Curriculum for Dental Education

Establishing Evaluation Criteria for Digital Dental Simulation Systems and the Differences between Software and Evaluators

J Dental Sci

emphasizes the importance of a "clinical clerkship" to guarantee the appropriateness and quality of students and to maintain patient safety and privacy [2]. Additionally, the revised version of this curriculum established simulation practice, model practice, and mutual practice, and provided learning objectives for skills education, including basic model practice before the start of the clinical clerkship.

Therefore, quantification of the evaluation of skills in basic model practice is required for dental students to treat patients in a clinical clerkship. It has been reported that teaching and evaluation of tooth preparation depend largely on the subjectivity of faculty staff or instructors for residents [3]. Dental students and clinical residents also understood that their image training was a major factor in their progress and they needed to observe faculty and supervisors as well as practice repeatedly [3]. Several reports on educational methods for cavity and tooth preparation and their evaluation focused on the angle (taper) of the axial surface and the amount of tooth removed in the tooth preparation for a full-cast crown [4-8]. There seems to be little emphasis on the development of specific educational methods to enable students and trainee dentists to perform ideal abutment tooth preparation and objective methods for evaluating the teeth after preparation [3]. Previous methods for evaluating tooth preparation have included mechanical measurements and, more recently, computer scanning [7-14]. Digital equipment has made it possible to obtain immediate data about the tooth preparation and provide effective feedback when teaching dental students about tooth preparation [13]. Therefore, it is expected that students using the simulation system to practice repeatedly will have their skills evaluated as numerical data, making the feedback more useful. At Kyushu Dental University, we have introduced the digital dental simulation system, which displays the formation status on the screen in real time. The hardware and software of the new device were developed in Taiwan and have already been introduced to dental schools and universities in Taiwan and throughout Asia. However, this is the first trial of the new device in Japan and its usefulness as an evaluation tool needs to be assessed. This preliminary study was conducted to confirm the relationship between the evaluation of tooth preparation by the new device and evaluation by a dentist to enable this system to be used as a new tool for quantitative evaluation of tooth preparation by students and residents in the future.

Materials and Methods

Subjects

This study was conducted with approval from the ethics committee of Kyushu Dental University (approval number: 18-49). A total of 187 students participated: 99 fifth-year students and 88 sixth-year students at the School of Dentistry, Faculty of Dentistry, Kyushu Dental University. The data was collected from 1 April 2019 to 31 March 2020.

Materials

SimEX DDS-100 (SimEX) (EPED Inc., Kaohsiung City, Taiwan)
Artificial teeth: A5A-500-#36 (mandibular left first molar, resin teeth) (Nissin Dental Products Inc., Kyoto, Japan)
Oral model: D16-500AQF (J. MORITA CORP., Osaka, Japan)
Bar: TR-13 (MANI Inc., Tochigi, Japan)
Dental turbine: PAR-4HEX O KV (J. MORITA CORP., Osaka, Japan)

Procedure

The maxillary and mandibular tooth molds were mounted on the phantom. The faculty staff or supervisor started the software including SimEX and performed a 3D calibration of the mandibular left first molar and turbine position. The subject used a dental turbine to create a full cast crown preparation on a non-vital mandibular left first molar on the model. The faculty staff or supervisor first evaluated the scores using SimEX. The evaluator then scored it according to the evaluation criteria.

Evaluation by software including SimEX

The software evaluated the teeth after preparation by dividing them into four cross-sections (old four cross-sections). The software was reprogrammed to re-evaluate the teeth in four cross-sections (new four cross-sections). The software was further programmed to reassess the teeth in eight cross-sections.

Evaluator

The evaluator was a dentist who is a faculty member of the Division of Clinical Clerkship Development and Research, Department of Oral Function, Kyushu Dental University.

Criteria for the evaluator

The following items were evaluated with a score of 2, 1, or 0: the amount of occlusal surface removed, the taper of the axial surface, the morphology of the chamfer-shaped margin, and the position of the margin (flush with the gingiva). The criteria are shown in Table 3.

Statistical Analysis

Statistical analysis was performed using JMP® 9.0.2 (SAS Institute Inc., Cary, NC, USA) and SPSS Statistics 25 (IBM, Armonk, NY, USA). A p-value of less than 0.05 was

Morishita M, et al. Establishing Evaluation Criteria for Digital Dental Simulation Systems and the Differences between Software and Evaluators. J Dental Sci 2021, 6(3): 000309.

Copyright© Morishita M, et al.

considered to indicate a statistically significant difference. Results are expressed as the mean ± standard deviation (SD). Spearman's rank correlation coefficient was conducted for each of the old four cross-sections, the new four cross-sections, and the eight cross-sections. A Wilcoxon rank-sum test was used to analyze the differences in each criterion of the evaluator. A multiple regression analysis was used to examine scores by the software and scores by the evaluator, with the objective variable being scores by the software from the eight cross-sections, and the explanatory variables being the scores, the scoring rate, and the dummy variables (0 or 1) for each criterion by the evaluator.

Results

First, we compared DentSim, which has been widely used, with SimEX (Table 1), which indicated that SimEX was more useful than DentSim in terms of applications, compatibility, and language support because SimEX is the more recent device.

	SimEX	**DentSim**
Manufacturer	EPED Inc.	Image Navigation Ltd.
Country of origin	Taiwan	Israel
In-house optical system	Yes	Yes
Application	Cavity, crown & bridge, endodontics, and pedodontics	Limited selection of lessons
Compatibility	Compatible with NISSIN, FRASACO, ACADENTAL, and KAVO tooth models	Only compatible with KAVO
Languages supported	5	3

Table 1: Comparison between SimEX and DentSim.

The results of the average scores by the software are shown in Figure 1. The software evaluated the amount of preparation and margin morphology on four cross-sections of the tooth and expressed them as a score. However, there was a large discrepancy between the visual evaluation of the formed teeth by faculty staff and the score evaluation by the software. This is because the scores by the software were high even for inappropriate forms. The average score of the old four cross-sections was very high, in the range of 80 points. Therefore, we asked the system developer to adjust the scores in the new four cross-sections and re-evaluated the scores using the same data. We also conducted an evaluation using eight cross-sections, hoping that increasing the number of reference sections would result in a more accurate evaluation than the four cross-sections. The average score of the evaluation for the new four and eight cross-sections was in the range of 60 points. No difference was observed in the average scores between the new four and eight cross-sections.

Figure 1: Average scores by the software according to the number of reference cross-sections.
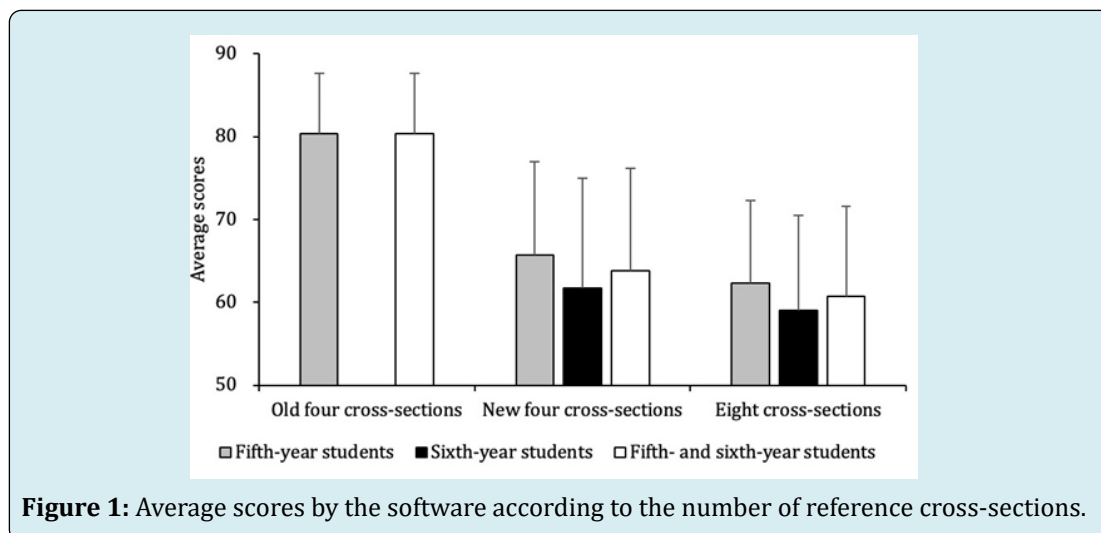
Table 2 shows the correlations between the scores of each of the old four, the new four, and the eight cross-sections. The results showed that the old four, new four, and eight cross-sections were statistically significantly correlated with each other. Based on these results, we decided to perform a software-based evaluation using eight cross-sections.

	Old four cross- sections	New four cross-sections	Eight cross-sections
Old four cross-sections	1.00	0.96***	0.93***
New four cross-sections	0.96***	1.00	0.95***
Eight cross-sections	0.93***	0.95***	1.00

***$p < 0.0001$

Table 2: Spearman's rank correlation coefficient associated with the average scores by the software according to the number of reference cross-sections.

Table 3 shows the criteria for the evaluator. The evaluator's results using these criteria showed that there were no statistically significant differences in each of the endpoints (Figure 2). The mean score rate was $31.2 \pm 16.3\%$, and there was no statistically significant difference among the subjects.

Scores	2	1	0
Amount of occlusal surface removed	Appropriate	Slightly too much or slightly too little	Obviously too much
Taper of the axial surface	No undercuts and a taper of less than 5 degrees	No undercuts and a taper of more than 5 degrees	Undercuts and a taper of more than 5 degrees
Morphology of the chamfer-shaped margin	Light chamfers all the way round the tooth	Deep chamfers around one-half of the tooth	No light chamfer all the way round
Position of the margin	Almost continuous with the gingival margin	Within 1 mm above and below the gingival margin and is continuous	Not continuous and significantly away from the gingival margin

Table 3: Criteria for the evaluator.

Table 4 shows the results of the multiple regression analysis of the scores by the software and the evaluator, where the objective variable was the score by the software in eight cross-sections, the explanatory variables were the score and the score rate by the evaluator, and a dummy variable (0 or 1) for each evaluation item. The results showed that the margin morphology evaluation by the evaluator was a significant explanatory variable for the score evaluation by the software.

	Unstandardized coefficients		Standardized coefficients		
	B	Standard error	Beta	t	Significance
(Constant)	54.275	2.495		21.576	0.000
Occlusal surface 1 point	-1.352	1.785	-0.062	-0.757	0.450
Occlusal surface 2 point	-1.044	2.579	-0.036	-0.405	0.686
Taper of the axial surface 1 point	2.157	2.407	0.099	0.896	0.371
Taper of the axial surface 2 point	2.242	2.612	0.096	0.859	0.392
Margin 1 point	8.050	2.204	0.368	3.653	0.000*
Margin 2 point	7.439	2.618	0.302	2.841	0.005*
Position 1 point	-1.600	1.847	-0.073	-0.866	0.388
Position 2 point	-0.616	2.255	-0.024	-0.273	0.785

R squared value=0.098, Adjusted R Squared value=0.057

Occlusal surface: Amount of occlusal surface removed; Margin: Morphology of the chamfer-shaped margin; Position: Position of the margin

Table 4: Multiple regression analysis between scores by the software and the evaluator.

Discussion

Many studies have used simulation software with evaluation systems such as DentSim and ClinSim [7-21]. A study comparing simulation using Simodont and actual clinical cases found that there was no difference and that the simulation system was an effective tool for education [14]. The new digital device, which was introduced to our university, can be incorporated into practical training because its initial calibration can be performed more quickly than DentSim, which makes it suitable for dental education (Table 1). Therefore, this study was conducted to examine the differences between evaluation by the new device software and a human evaluator. In the initial version of the software, four cross-sections were used as reference cross-sections for evaluation (old four cross-sections), and as shown in Figure 1, the score was 80.4 ± 7.3. It was clear that the standard of tooth preparation that the human evaluators considered to be inappropriate, such as undercutting, excessive cutting, and extremely poor tooth preparation, was evaluated with a high score by the software. We therefore decided to recalibrate the software and to increase the number of reference cross-sections to make it more precise. We asked the system

developer to change the following settings:
(1) new evaluation using four cross-sections (new four cross-sections)
(2) new evaluation using eight cross-sections (eight cross-sections).

After the program was modified, we re-evaluated our previous data with the new four and eight cross-sections. The average scores were 63.8 ± 12.4 and 60.8 ± 10.9, respectively. As shown in Table 2, it is clear that the scores in the old four, new four, and eight cross-sections were significantly correlated with each other. Because the scores by the software and by the human evaluator were generally close, we used the scores from the eight cross-sections.

Figure 2 shows the results of the evaluation of the artificial tooth after preparation by a human evaluator using the evaluation criteria shown in Table 3. Although the scores by the software and the human evaluator were not compared according to the same criteria, the software scores tended to be higher than the human evaluator scores.
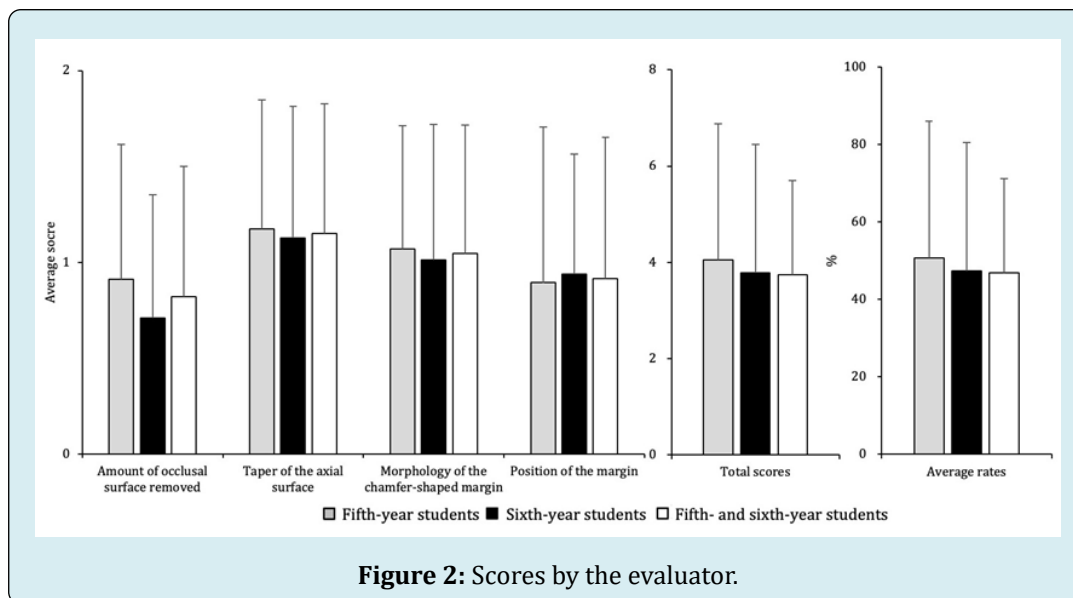
Figure 2: Scores by the evaluator.

Taguchi, et al. also compared the results of subjective and objective evaluations in the same way, and although there was a correlation between the two results, they pointed out that questions remained [3]. To clarify the relationship between the software scores and the criteria for the scores by the evaluator, we conducted a multiple regression analysis using dummy variables, with the software scores as the objective variable and each of the criteria items for the evaluator as shown in Table 3 as explanatory variables. The results are shown in Table 3, and it is clear that the scores for the morphology of the chamfer-shaped margin by the

human evaluator are a significant explanatory variable for the software scores. Although the human evaluator scored the items according to the criteria, the evaluation was not objective (i.e., measuring instruments were not used) as in other studies, and variations due to the subjective evaluation may have affected the results. One of the evaluation items, the amount of occlusal surface removed, defined a score of 0 as being "obviously too much", but the definition of "obviously" is not clear. Such ambiguous criteria require calibration when there are multiple evaluators, but because only one evaluator scored the artificial teeth in the present

study, the lack of clarity was considered to be minimal. The morphology of the chamfer-shaped margin was considered to be a significant explanatory variable because it assessed whether the light chamfer-shape covered the whole tooth or more than half of the tooth; this may have been easier to assess both visually and subjectively than the other criteria. This study was the first to compare the use of new device with a human evaluator. Our results revealed that the scores by the human evaluator were significantly related to the scores by the software. Limitations of this study were that it was conducted by a single evaluator, and the results were based on subjective evaluation rather than evaluation using measurement instruments. If multiple raters had been used, it remains unclear how the evaluators' scores would correlate with the scores by the software. Additionally, we evaluated tooth preparation only in the mandibular left first molar, so we could not examine the effect of tooth location. In future research, we aim to divide the evaluation items into those related to the amount of tooth removed and those related to the shape, and to examine the relationship between the scores by software and the scores by the human evaluators, dividing the groups into those with high scores and low scores. We would also like to demonstrate that the new digital device is a useful dental education tool for undergraduate students and residents.

Conclusion

Our work demonstrates for the first time that the scores by the human evaluator are significantly related to the scores by the software using SimEX, a new digital dental education device with a real-time evaluation function. It shows that SimEX will be able to provide a quantitive evaluation of tooth preparation. Our findings suggest that SimEX may be helpful as a new digital dental educational tool.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgments

We thank Helen Jeays, BDSc AE, from Edanz Group (https://en-author-services.edanz.com/ac) for editing a draft of this manuscript.

References

1. Miyazaki T, Hotta Y (2011) CAD/CAM systems available for the fabrication of crown and bridge restorations. Aust Dent J 56(1): 97-106.

2. The Ministry of Health, Labor and Welfare (2016) The model core curriculum for dental education in Japan [Only in Japanese].

3. Taguchi Y, Takizawa T, Oka Y, Fujii N (2012) A Study of Clinical Training and Evaluation for the Preparation of a Full-Veneer Crown. Ann Jpn Prosthodont Soc 4: 434-443.

4. Kanuma A, Kashiwada T, Kimura K, Yoshida K (1974) A Report on the Preparation of the Maxillary Left First Molar for a Full Cast Crown in Manikin Head. J Jpn Prosthodont Soc 18(1): 70-84.

5. Kimura K, Endo Y, Ishibashi M, Kasahara S, Hatakeyama N, et al. (1991) A study on The Preparatory Result of The Maxillary Left First Molar for a Full Cast Crown in Manikin Head before and after Policlinical Practices. J Jpn Prosthodont Soc 35(1): 125-136.

6. Nameta K, Shimizu K, Chiba O, Mikami J, Sasagawa T, et al. (1992) A Study on the Tooth Preparation for Full Coverage Crown. Part 1. Evaluation of the form of abutment teeth prepared under clinical situation. J Jpn Prosthodont Soc 36(1): 181-195.

7. Yamaguchi T, Hosoya N, Kurachi Y, Yoshida T, Morito A, Koda S, Yamamoto T, Momoi Y (2012) Objective Evaluation of Access Cavity Preparation. Jpn Soc Conserv Dent 55(4): 278-284.

8. Wierinck E, Puttemans V, Swinnen S, van Steenberghe D (2005) Effect of augmented visual feedback from a virtual reality simulation system on manual dexterity training. Eur J Dent Educ 9(1): 10-16.

9. Quinn F, Keogh P, McDonald A, Hussey D (2003) A study comparing the effectiveness of conventional training and virtual reality simulation in the skills acquisition of junior dental students. Eur J Dent Educ 7(4): 164-169.

10. Okuyama Y, Kasahara S, Kimura K (2005) Quantitative evaluation of axial wall taper in prepared artificial teeth. J Oral Sci 47(3): 129-133.

11. Ohyama A, Araki K, Nitta H, Shimizu C, Ohara S, et al. (2007) The roles of clinical skills laboratory in dental education. Health Sci and Health Care 7(2): 78-82.

12. Welk A, Maggio MP, Simon JF, Scarbecz M, Harrison JA, et al. (2008) Computer-assisted Learning and Simulation Lab with 40 DentSim Units. Int J Comput Dent 11(1): 17-40.

13. Abe T, Okuyama Y, Kasahara S, Kimura K (2009) Educational Effect on Tooth Preparation of Visual Feedback using Computer Graphics. Ann Jpn Prosthodont Soc 1: 123-129.

14. Mirghani I, Mushtaq F, Allsop MJ, Al Saud LM, Tickhill

N, et al. (2018) Capturing differences in dental training using a virtual reality simulator. Eur J Dent Educ 22(1): 67-71.

15. Quinn F, Keogh P, McDonald A, Hussey D (2003) A pilot study comparing the effectiveness of conventional training and virtual reality simulation in the skills acquisition of junior dental students. Eur J Dent Educ 7(1): 13-19.

16. Welk A, Splieth C, Rosin M, Kordass B, Meyer G (2004) DentSim – A Future Teaching Option for Dentists. Int J Comput Dent 7(2): 123-130.

17. Persson A, Andersson M, Oden A, Englund GS (2006) A three-dimensional evaluation of a laser scanner and a touch-probe scanner. J Prosthet Dent 95(3): 194-200.

18. Hollis W, Darnell LA, Hottel TL (2011) Computer Assisted Learning: A New Paradigm in Dental Education. J Tenn Dent Assoc 91(4): 14-18.

19. Yamada S, Ono Y, Sugiyama T, Watarai M, Asari J, et al. (2015) Method of Evaluation of Preclinical Tooth Modeling Practice in Pediatric Dentistry and Automatic Evaluation of 3D Shape. Jpn J Pediatr Dent 53(4): 471-477.

20. Nakagawa S, Okawa T, Ito Y, Fukumoto T, Wu H, et al. (2016) Evaluation the AR Navigation System for preparation of the facings of hard-resin abutment teeth. J Osaka Odontol Soc 79(1): 17-22.

21. Dixon J, Towers A, Martin N, Field J (2021) Re-defining the virtual reality dental simulator: Demonstrating concurrent validity of clinically relevant assessment and feedback. Eur J Dent Educ 25(1): 108-116.