

RESEARCH

Open Access



Video-enhanced training for entrustable professional activities in preclinical oral surgery: a randomized controlled trial

Yue Lin^{1,2†}, Zhe Wen^{3†}, Yuqi Li^{4†}, Xiaosong Wei¹, Jiafen Zhang¹, Tinglin Yan^{1*} and Zhexuan Bao^{1*}

Abstract

Background Traditional preclinical teaching methods face challenges in adequately preparing oral and maxillofacial surgery (OMFS) students for the entrustable professional activities (EPA) required in their future clinical internships. This study aimed to evaluate how a 4 K live-video teaching model (LTM), integrated into a preclinical transitional course (PTC), improves students' EPA-related performance.

Methods Thirty-six fifth-year undergraduate students participated in a PTC. They were randomly divided into two groups. The control group was taught using a traditional teaching model within the PTC (direct guidance at a manikin), while the experimental group received instruction supplemented with the LTM. The module incorporated high-fidelity simulation, wherein a physician acted as a simulated patient to provide real-time verbal feedback. After the PTC, students' competencies were comprehensively evaluated across several domains: comprehensive EPA scores (self-assessed, teacher-rated, and composite), procedural skills assessment scores, theoretical examination scores, and questionnaire-based self-evaluations and teacher evaluations of student outcomes. All assessment instruments were found to be reliable, as indicated by their ICC and Cronbach's α values.

Results The assessment instruments showed high validity. Compared to the control group, the experimental group achieved significantly higher scores in EPA self-assessment (93.05 ± 1.26 vs. 89.21 ± 1.26), teacher evaluation (88.79 ± 0.85 vs. 85.86 ± 0.85), and overall EPA performance (90.92 ± 0.91 vs. 87.54 ± 0.91) (all $p < 0.05$). The experimental group also demonstrated superior procedural skill test scores (89.49 ± 0.73 vs. 79.48 ± 0.73 , $p < 0.001$), self-evaluations of learning (21.61 ± 1.14 vs. 13.05 ± 1.70 , $p < 0.001$), and teacher evaluations of their performance (21.37 ± 1.23 vs. 18.93 ± 1.31 , $p < 0.001$). No significant difference was found in theoretical examination scores ($p > 0.05$).

Conclusion In this study, the integration of LTM into the OMFS PTC was associated with a significant enhancement in students' foundational procedural skills, self-perceived competence, and overall performance as evaluated by faculty. These findings suggest that this teaching model can be an effective tool for improving EPA-related performance in

[†]Yue Lin, Zhe Wen and Yuqi Li contributed equally to this work.

*Correspondence:

Tinglin Yan
253685448@qq.com
Zhexuan Bao
baozhexuan@163.com

Full list of author information is available at the end of the article



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

a preclinical setting. However, given the single-center design and small sample size, further research is warranted to validate these results. This model shows potential for better preparing students for the clinical entrustment process.

Keywords Entrustable professional activities, Preclinical education, Oral and maxillofacial surgery, Dental education, Live video teaching

Introduction

Oral and maxillofacial surgery (OMFS) is closely linked with general surgery, otolaryngology, ophthalmology, anesthesiology, pediatrics, and other clinical departments. As a result, most treatments in OMFS involve invasive procedures. Key aspects in preparing students for the clinical practice of this field include understanding various maxillofacial diseases, techniques for doctor–patient communication, and various classic surgical methods, as well as the management of potential intra-operative and postoperative complications [1]. The traditional preclinical teaching model, however, often struggles to provide students with a comprehensive mastery of the required foundational skills, creating an urgent need for optimization. This educational challenge is not unique to OMFS but reflects a widespread issue across medicine that has boosted significant reforms since the late 20th century [2]. A key outcome of these reforms is Competency-Based Medical Education (CBME), an outcome-oriented approach focused on creating clinically competent physicians [3].

Within the CBME framework, Entrustable Professional Activities (EPAs) have emerged as the critical tool for translating competencies into observable clinical tasks. However, the integration of EPAs into dental education has been a recent and unevenly paced development [4, 5]. Key challenges persist globally, including defining specialty-specific tasks, ensuring consistent assessment, and providing adequate faculty development [6]. A significant gap remains in preparing students for high-stakes specialties like OMFS before they enter the clinical environment.

To understand why traditional teaching methods may hinder learning, it is useful to apply cognitive load theory (CLT). CLT posits that learning is most effective when instructional design manages the limitations of human working memory. The cognitive load consists of three components: (1) Intrinsic Load, which is the inherent difficulty of the subject matter itself (e.g., the complex steps of a tooth extraction); (2) Extraneous Load, which is the “bad” load generated by instructional design (e.g., a poor viewing angle or confusing explanations); (3) Germane Load, the load related to the process of deep learning and schema formation. Therefore, the primary goal of instructional design is to minimize extraneous load to free up cognitive resources that can be devoted to managing intrinsic load and promoting germane load [7].

Applying this CLT framework to OMFS preclinical education reveals a critical flaw in the traditional model. Procedures like tooth extraction are inherently complex, imposing a high intrinsic load due to their numerous, interacting steps [8]. Traditional teaching methods exacerbate this challenge by introducing an extraneous load. Sources of this load include obstructed viewing angles that force mental reconstruction, non-repeatable demonstrations that hinder reinforcement, and abstract verbal explanations that lack concrete visual grounding [7]. This high extraneous load consumes valuable cognitive resources that should instead be dedicated to germane load, that is, the deep processing required to build lasting mental schemas for the procedure. Consequently, students may fail to develop the robust foundational skills necessary for future clinical reasoning [9]. Real-time video technology directly targets these shortcomings. By providing a multi-angle and unobstructed view, it can transform abstract procedural demonstrations into concrete learning experiences, therefore minimizing this extraneous cognitive load and performing better on procedural tasks.

Despite this clear potential, previous studies on real-time video teaching in OMFS have primarily focused on enhancing theoretical knowledge or basic procedural skills, while largely overlooking its value for developing EPAs in a preclinical context. While clinical studentships are the ultimate setting for assessing EPAs, robust preclinical training and assessment are crucial for ensuring student readiness and patient safety. Therefore, this study introduces a 4 K live video teaching model (LTM) into a preclinical transitional course (PTC) for OMFS students to examine its impact on students’ performance on EPAs [10]. This study is the first to embed real-time video teaching within an EPA framework in a preclinical simulation setting, systematically evaluating its impact on EPA-based competency-based instruction.

Many studies have explored video-assisted surgical training. However, a critical gap remains, since few have integrated this technology with a competency-based framework like EPAs in a preclinical setting. Our study addresses this gap with two key contributions. First, we combined an LTM with an EPA assessment framework in a preclinical OMFS course. This approach shifts the educational focus from mastering several isolated skills toward achieving demonstrable readiness for entrustment. Second, we developed and validated a customized EPA scale for this purpose, providing a tool to assess

competence in a simulated environment. Therefore, our primary objective is to test the hypothesis that this integrated teaching model significantly enhances students' EPA-related performance, strengthening their foundational preparedness for future clinical entrustments.

Study design and methods

General background and data

The study took place in the Preclinical Simulation and Training Center, which is housed within the hospital's Comprehensive Department II, a general teaching department responsible for undergraduate dental education. This study was approved by the Affiliated Stomatological Hospital of Guangzhou Medical University Ethics Committee, with a registration number of LCYJ20250618007. From June to August 2025, the department recruited 36 students who were due to start

their internships in September, comprising 19 males and 17 females. Since the research involved no real treatments, a clinical trial number was not required.

The core of the study was a simulation-based PTC. Before the start of this PTC, all participants attended a 6-hour preparatory course on the foundational theoretical knowledge of OMFS. Once the course was over, they underwent a basic knowledge assessment (Fig. 1). There were no statistically significant differences between the two groups in terms of gender, age, or pre-PTC professional assessment scores. Student allocation was conducted by a third-party administrative staff member from the finance department who had no contact with the students or knowledge of the study. A blocked randomization sequence with a block size of four was generated using a computer-based random number generator (sealedenvelope.com, Sealed

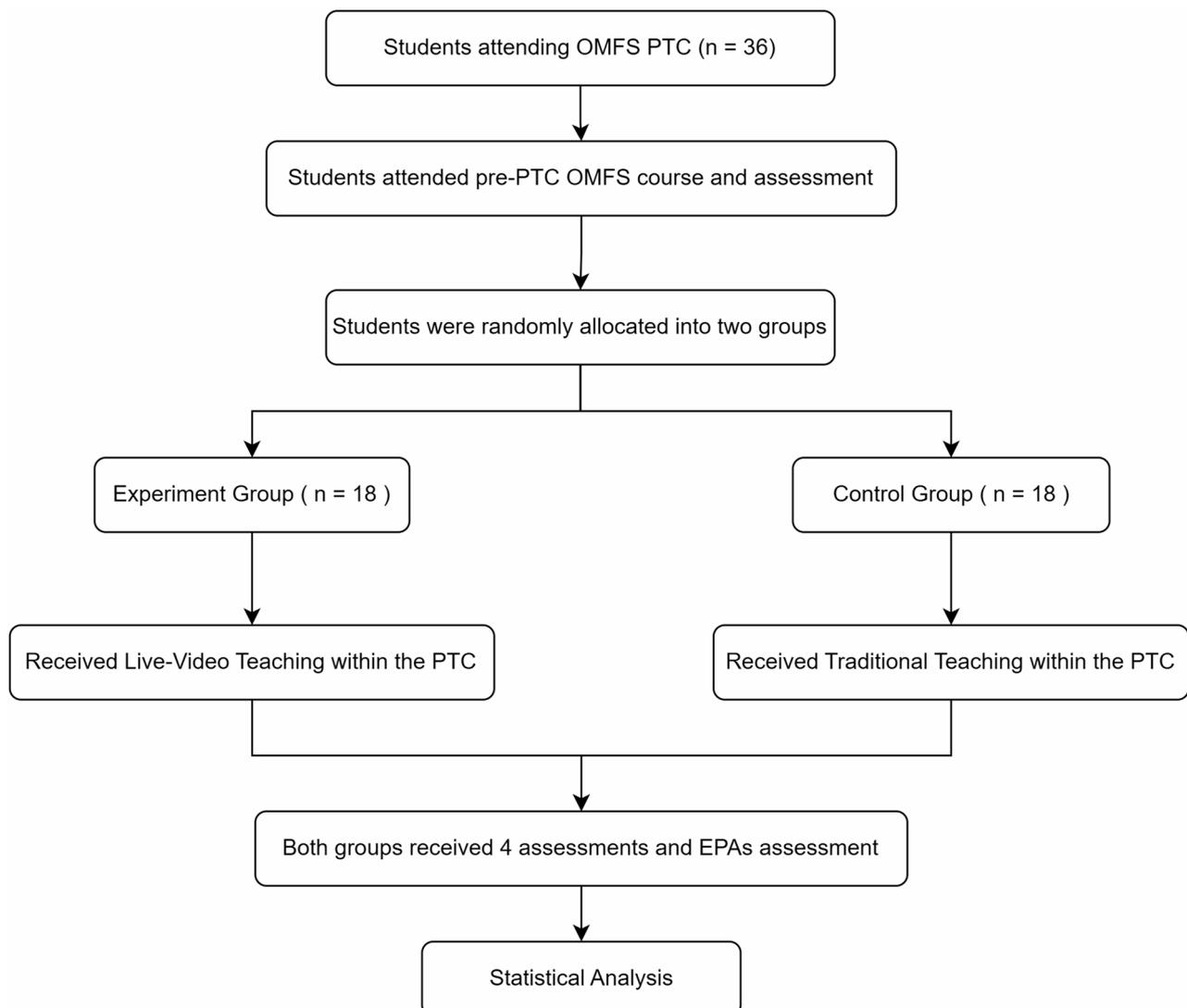


Fig. 1 Study flowchart

Envelope Ltd., London, U.K.) to ensure a 1:1 allocation ratio. To ensure allocation concealment, the third party prepared 36 sequentially numbered, opaque, sealed envelopes (SNOSE), each containing the group assignment for a single participant. After a student provided written informed consent, the lead instructor of the study contacted the third party, who then provided the next consecutively numbered envelope. The envelope was opened in the presence of the student to reveal their group assignment.

Experimental methods

The PTC framework and general procedures

Both groups were trained using the simulation-based PTC. All structured training sessions were held from Monday to Friday, with each student participating for 3 h per day. Throughout the PTC, three attending physicians rotated as both instructors and simulated patients. They were assigned to teach sessions based on a randomly designed schedule, ensuring that each instructor conducted sessions in both the control and experimental groups. They presented all students with scripted case scenarios that described specific diseases and symptoms. The instructors also performed procedures on a manikin (SB23464, Nasco Healthcare, Saugerties, NY, USA). All sessions, regardless of group, followed a standardized SOP comprising case presentation, instructor demonstration, student hands-on practice on the manikins, and post-session debriefing. This rotation system was implemented to mitigate potential instructor bias by balancing the influence of individual teaching styles across both groups.

The training required students in both groups to demonstrate competence across six key domains, which consisted of: (1) doctor–patient communication: gaining simulated patients' initial trust through effective and clear communication; (2) medical history collection: accurately and concisely collecting information on simulated patients' chief complaints, present illnesses, and past medical histories; (3) physical examination: conducting comprehensive and detailed examinations and accurately identifying urgent issues on the manikin; (4) operational skills: performing further practices after examinations on the manikin; (5) health education: providing

simulated patients with postoperative instructions; (6) humanistic care: emphasizing simulated patients' safety and risk management throughout the encounter. Written informed consent was obtained from all participating students and instructors before the module began. These domains were designed under the framework of the American Dental Education Association (ADEA)'s 'Competencies for the New General Dentist', adapted for our specific preclinical context [11].

Control group

Students in the control group were trained using a traditional teaching model within the PTC (referred to as the traditional-teaching-model PTC, the TTM). In this model, instructors provided direct, at-the-chairside supervision and real-time verbal guidance as the students performed procedures on the manikins.

Experimental group

The experimental group employed an improved teaching model that supplemented the teaching model of the PTC with live video instruction (hereafter referred to as the 4 K Live video Teaching Model, LTM). This involved real-time teaching using a 4 K miniature dental camera, with all key procedures recorded in real-time and archived on local and cloud-based storage (Fig. 2). In addition to the instructor's live explanations during procedures, students observed the operations in real-time on monitors and participated in post-session debriefing sessions using the video recordings for further review and discussion. The 4 K imaging system was permanently installed in the same preclinical simulation and training center and was switched on or off solely according to the daily roster.

To ensure the reproducibility and consistency of the intervention, a standardized protocol was implemented for the setup and use of the LTM. The miniature 4 K camera was mounted directly onto the central handle of the instructor's dental operating light. This fixed position ensured a consistent, co-axial perspective that precisely mimicked the instructor's primary line of sight for all procedures.

As the camera was coupled with the operating light, the viewing angle and illumination were inherently standardized. The instructor would naturally direct the



Fig. 2 Schematic Illustration of the Live Video Teaching Model

light onto the operative field, which was the manikin's oral cavity. This setup guaranteed that the entire operative field was always optimally lit and captured ideally. The live feed was broadcast to 18 24-inch HD monitors at each student station. The monitors were positioned at eye level, approximately 60–70 cm from the student, to provide a clear and ergonomic viewing experience. All instructional demonstrations were recorded in their entirety. Recordings were systematically archived using a standardized naming convention (YYYY-MM-DD_ProcedureName). During post-session discussion, these videos were used for detailed review, allowing instructors to pause and replay critical steps for discussion.

Assessment criteria

Multi-dimensional assessment criteria

Procedural skills assessment tools Upon completing their PTC before leaving the department, students were assessed on their comprehensive procedural skills in a simulated environment. The assessment covered three main areas, and was adapted from the Chinese National Dental Practitioner Examination: (1) preoperative preparations for common maxillofacial diseases in a simulated context: this included doctor–patient communication, understanding indications and contraindications, and preparing instruments; (2) procedural operations on a manikin: this included local anesthesia and tooth extraction procedures; (3) post-procedural tasks: this included postoperative care instructions and medical record documentation.

Three senior attending physicians from the other campus, who were unaware of this experiment, served as independent raters throughout the whole experiment, scoring each student based on the Chinese National Dental Practitioner Examination rating instructions, aiming to evaluate their basic capabilities in diagnosing, treating, and preventing common and frequently occurring diseases in OMFS. The total score was 100 points.

Theoretical comprehensive assessment At the end of the PTC, teaching instructors conducted a comprehensive theoretical assessment of the students. The content was jointly designed by three other senior attending physicians and three other teachers from the Department of Oral and Maxillofacial Surgery, and covered basic theoretical knowledge, clinical diagnosis, treatment, and emergency handling principles and methods of treating various common maxillofacial diseases, as well as the

interpretation of imaging examinations for common oral diseases. Students were scored by a teacher from the other campus with a total score of 100 points.

Self-Evaluation of learning outcomes After the PTC, the lead instructor administered an anonymous questionnaire survey to evaluate each student's self-assessment of their learning outcomes. The questionnaire included five aspects: (1) stimulating interest in learning, (2) improving logical thinking ability, (3) enhancing procedural skills, (4) improving doctor–patient communication skills in a simulated environment, and (5) enhancing mutual cooperation abilities. Each item was rated on a Likert scale with five levels: 1 for “strongly disagree”, 2 for “disagree”, 3 for “neutral”, 4 for “agree” and 5 for “strongly agree”, with a maximum score of 25 points. This questionnaire was designed specifically for this research. It is available as Supplementary File 1.

Teacher evaluation of student PTC outcomes After the PTC, the comprehensive performance of each student during the PTC was evaluated by the three teaching physicians. The evaluation covered five aspects: (1) simulated doctor–patient communication; (2) aseptic concept; (3) case selection/anesthesia technique; (4) operational skills; (5) medical record writing. Each item was rated on a Likert scale with five levels: 1 for “very poor”, 2 for “poor”, 3 for “neutral”, 4 for “good” and 5 for “very good”, with a maximum score of 25 points. This questionnaire was also created only for this research. It is available as Supplementary File 2.

EPA assessment scoring

Based on the characteristics of undergraduate students in PTC, this study developed detailed EPA rubrics, including self-assessment scores, teacher assessment scores, and comprehensive scores, each with a maximum score of 100 points (Tables 1 and 2). As self-evaluation and faculty evaluation are of equal importance in EPA assessment, each contributes 50% to the final score.

The EPA Scale was developed using key points from the National Licensing Examination, the Attending-Physician Qualification Examination, and standard oral–oral–maxillofacial surgery textbooks. Eight oral–maxillofacial surgery experts with ≥ 10 years of teaching experience (four professors, four associate professors) took part in a two-round Delphi review. In round one the experts rated 30 items on a four-point relevance scale, then 14 items with an I-CVI < 0.78 were revised and merged into seven. After round two, the final 23 items achieved an I-CVI of 0.81–0.94 and an S-CVI/Ave of 0.94, meeting the threshold proposed by Polit and Beck and indicating acceptable content validity for the checklist. The EPA Assessment

Table 1 EPA assessment scoring summary

Assessment Category	Weight	Score	Total Score
Self-Evaluation	50%		
Teacher Evaluation	50%		

Table 2 EPAs scoring details

First-Level Indicator	Second-Level Indicator	Evaluation Criteria	Weight	Self-Evaluation	Teacher Evaluation
1. Operation Process – 80 Points	1.1 Preoperative Preparation for Tooth Extraction - Concept of Patient Care, Preoperative Evaluation 1	All maneuvers must be executed gently yet efficiently, prioritizing the simulated patient's comfort and minimizing trauma.	3%	50%	50%
	1.2 Preoperative Preparation for Tooth Extraction - Concept of Patient Care and Preoperative Evaluation 2	Confirm the identity of the tooth scheduled for extraction and perform a systematic intra-oral examination of the operative site.	4%	50%	50%
	1.3 Preoperative Evaluation: Positioning 1	Manikin Positioning: 1. Maxillary extractions: align the maxillary occlusal plane at $\approx 45^\circ$ to the floor, so the operative field lies between the surgeon's shoulder and elbow level. 2. Mandibular extractions: maintain the mandibular occlusal plane parallel to the floor, positioning the operative field just below elbow height.	3%	50%	50%
	1.4 Preoperative Evaluation: Positioning 2	Surgeon's Positioning: The surgeon should stand anterolateral to the manikin's right side, maintaining a balanced, ergonomically neutral posture with the upper body fully relaxed.	3%	50%	50%
	1.5 Preoperative Evaluation: Positioning 3	Light's positioning: Adjust the operating light centrally over the field to offer shadow-free illumination of the operative site.	2%	50%	50%
	1.6 Preoperative Evaluation: Instrument Check 1	Confirm availability of an oral examination tray, mucosal antiseptic, cotton swabs/balls/rolls, and all required dressings.	2%	50%	50%
	1.7 Preoperative Evaluation: Instrument Check 2	Ensure that gingival elevators and periodontal curettes are present, sterile, and able to be used.	2%	50%	50%
	1.8 Preoperative Evaluation: Instrument Check 3	Select and verify the sterility of the appropriate dental elevators and extraction forceps.	3%	50%	50%
	1.9 Basic Tooth Extraction Steps - Gingival Separation 1	Re-identify the tooth and disinfect the gingival tissue of the affected tooth.	4%	50%	50%

Table 2 (continued)

First-Level Indicator	Second-Level Indicator	Evaluation Criteria	Weight	Self-Evaluation	Teacher Evaluation
1. Operation Process – 80 Points	1.10 Basic Tooth Extraction Steps - Gingival Separation 2	Use the gingival separator to fully separate the gingiva, correctly utilizing the working surface of the separator, and ensure proper use of a fulcrum.	5%	50%	50%
	1.11 Basic Tooth Extraction Steps - Use of Dental Elevator 1	Reconfirm the tooth scheduled for elevation.	3%	50%	50%
	1.12 Basic Tooth Extraction Steps - Use of Dental Elevator 2	Loosen the affected tooth with a dental elevator: The dental elevator can be held in a palm-and-thumb or finger-rest grasp. Insert horizontally (using the alveolar crest as a fulcrum) or vertically (using the interdental septum as a fulcrum) while the non-dominant left hand should simultaneously support the affected tooth and the adjacent tooth.	8%	50%	50%
	1.13 Basic Tooth Extraction Steps - Use of Extraction Forceps 1	Holding the extraction forceps: Place the forceps handle in the palm, gripping one handle with index-to-little fingers, while the other handle rests against the palm. Meanwhile, stabilize the hinge with the thumb for maximal control.	4%	50%	50%
	1.14 Basic Tooth Extraction Steps - Use of Extraction Forceps 2	Placement of Extraction Forceps: Seat the beaks parallel to the tooth's long axis, advancing apically until they fully engage the root surface.	7%	50%	50%
	1.15 Basic Tooth Extraction Steps - Use of Extraction Forceps 3	Tooth Luxation: Based on the anatomical characteristics of the tooth roots, correctly use the combination of rotational, rocking, and pulling forces to luxate the affected tooth along the path of least resistance.	7%	50%	50%
	1.16 Basic Tooth Extraction Steps - Use of Extraction Forceps 4	Continuously protect the marginal gingiva and adjacent teeth from inadvertent trauma during forceps manipulation.	4%	50%	50%
	1.17 Post-Extraction Management 1	Examine the extracted tooth to confirm root integrity and verify that the number of roots matches expected anatomy.	4%	50%	50%
	1.18 Post-Extraction Management 2	Gently probe the socket with a surgical curette to remove debris or simulated calculus, avoiding aggressive scraping.	4%	50%	50%
	1.19 Post-Extraction Management 3	Compress the socket walls with a sterile cotton roll or gauze to re-approximate the alveolar bone.	4%	50%	50%
1.20 Post-Extraction Management 4	Performing hemostasis procedure by applying firm pressure with a sterile cotton roll or gauze pad.	4%	50%	50%	
2. Operation Outcome – 20 Points	2.1 Tooth Extraction Outcome 1	Extraction completed within the allotted operation time.	10%	50%	50%
	2.2 Tooth Extraction Outcome 2	The extracted tooth exhibits intact, undamaged root(s).	5%	50%	50%
	2.3 Tooth Extraction Outcome 3	No iatrogenic injury to opposing dentition, alveolar bone, or marginal gingiva.	5%	50%	50%

Scoring was conducted by the previously mentioned three senior attending physicians from the other campus.

To ensure the robustness of the newly developed OMFS-EPA scale, we assessed several key psychometric properties.

Content Validity was established through a two-round Delphi method with eight OMFS experts, as previously described. The final 23 items achieved a satisfactory S-CVI/Ave of 0.94, indicating that the scale's content is highly relevant to the preclinical OMFS training context.

Evidence for construct validity was gathered by examining the relationships between EPA scores and other

measured outcomes. Specifically, we hypothesized that our EPA scale would show convergent validity by positively correlating with the objective procedural skills assessment scores, as both aim to measure practical competence. Conversely, we expected it to demonstrate discriminant validity by showing a weaker correlation with the theoretical comprehensive assessment scores, as the scale is designed to evaluate applied skills rather than declarative knowledge. The results of these analyses, presented in the Results section, support these hypotheses.

Internal consistency for both the self-assessment and teacher-evaluation components of the EPA scale was planned to be assessed using Cronbach's α coefficient.

Inter-rater reliability (IRR) for the teacher evaluation component was a critical focus. Prior to the formal assessment, the three external senior attending physicians serving as raters participated in a two-hour calibration session. During this session, they reviewed the EPA rubric in detail and collectively scored two standardized video-recorded performances to align their scoring standards. The IRR was subsequently calculated using a two-way mixed-effects intraclass correlation coefficient (ICC) model with an absolute agreement definition. This specific model (ICC 3,k) was chosen because our three raters were fixed and we were interested in the reliability of their averaged ratings. The pre-assessment video-based resulting ICC was 0.86, which is considered to indicate excellent agreement among the raters.

Assessment bias

To minimize assessment bias, a strict blinding protocol was enforced for all external evaluators. All participating students were assigned a unique, non-identifying participant code (e.g., Student 01, Student 02) at the beginning of the study. All assessment materials were labeled only with these codes. The evaluators were never informed of the students' group allocation, and their only task was to score. The only exception to this protocol was the 'Teacher Evaluation of Student PTC Outcomes,' an outcome which was rated by the unblinded teaching physicians to capture their direct pedagogical impressions. As for missing data, all 36 enrolled participants completed the entire study protocol and all post-intervention assessments. Consequently, there were no missing data in this study.

Statistical methods

The assessment scores of students from both groups were analyzed using IBM SPSS Statistics (Version 22.0; IBM Corp., Armonk, NY, USA). A p -value of <0.05 was considered statistically significant.

Results

Reliability analyses

After scoring was completed, reliability analyses were conducted. The consistency of the assessments was evaluated using Cronbach's α coefficients. Inter-rater reliability

for assessments involving multiple raters was calculated using the intraclass correlation coefficient (ICC) based on a two-way mixed-effects model. For the procedural skills assessment, the Cronbach's α was 0.91, and the ICC (3,3) was 0.85. For the self-evaluation of learning outcomes, the Cronbach's α was 0.89. Meanwhile, the teacher evaluations of student outcomes yielded a Cronbach's α of 0.88 and an ICC (3,3) of 0.82. For the EPAs, the self-assessment Cronbach's α was 0.91, the teacher-evaluation Cronbach's α was 0.92, and the ICC (3,3) was 0.87.

Comparison of general conditions

This study enrolled 36 students with a mean age of 23.56 ± 0.94 years; 19 were male (52.78%) and 17 female (47.22%). The experimental group ($n=18$) had a mean age of 23.83 ± 1.10 years and comprised 11 males (61.11%) and seven females (38.89%); the control group ($n=18$) had a mean age of 23.28 ± 0.67 years, with eight males (44.44%) and 10 females (55.56%). As data for age was not normally distributed ($p < 0.001$ for both groups on age), the Mann-Whitney U test was used for the comparisons. Results indicated no significant difference between the groups in age ($U=120.5$, $p=0.133$). A chi-square test of independence was used to analyze the distribution of sex, and an independent-samples t-test was used to compare pre-PTC professional assessment scores between the groups, as they were both normally distributed. No significant differences were found between groups in age, sex, or pre-PTC professional assessment scores ($p=0.133$, $p=0.317$, $p=0.651$, respectively). Detailed values are presented in Table 3.

Comparison of theoretical comprehensive assessment and procedural skills assessment scores between the two groups

At the end of the PTC, the theoretical comprehensive assessment and the procedural skills assessment were conducted separately for the two groups of students. For the theoretical comprehensive assessment, the Shapiro-Wilk test yielded $p=0.489$ for the experimental group and $p=0.064$ for the control group, while Levene's test produced $p=0.987$. Likewise, after Shapiro-Wilk testing (experimental group $p=0.105$; control group $p=0.226$) and Levene's test ($p=0.192$). To analyze the group differences on the procedural skills assessment scores and the theoretical comprehensive assessment, analysis of covariance (ANCOVA) was used to control the potential confounding effect of the covariate (pre-PTC professional assessment scores). The data for both dependent variables met the assumption of homogeneity of regression slopes ($p=0.477$ for procedural skills assessment scores and $p=0.878$ for theoretical comprehensive assessment, respectively), and the covariate did not have a statistically

Table 3 Comparison of gender between experimental and control groups

Group	Male	Female
Experimental ($n=18$)	11	7
Control ($n=18$)	8	10
p -value	0.317 ($p > 0.05$)	

significant effect in either of the main models ($p = 0.571$ and $p = 0.055$, respectively).

The theoretical comprehensive assessment scores for the experimental group and the control group indicated no significant difference in theoretical learning between the two groups. The procedural skills assessment scores for the two groups showed a significant difference in procedural skills learning. This suggests that there was no significant difference in the impact of the two teaching models on students' theoretical learning; however, live video teaching could significantly improve the operational skills of students for future internship performance. Specific data are shown in Table 4.

Comparison of Self-Evaluation of learning outcomes between the two groups of students

A survey was designed to understand the self-evaluation of learning outcomes by the two groups of students. Table 5 shows the specific content and statistical results of the survey. The hypothesis that the scores were normally distributed was rejected as the data were not normally distributed, confirmed by the Shapiro–Wilk test ($p < 0.05$), therefore, the Mann–Whitney U test was used. The results indicated that the self-evaluation scores of the experimental group were significantly higher than those of the control group in all five aspects: stimulating interest in learning, improving logical thinking ability, enhancing operational skills, improving doctor–patient communication skills, and enhancing mutual cooperation abilities, as well as overall satisfaction. There were significant differences between the two groups, indicating that live video teaching could boost the students' self-perceived learning outcomes. Specific data are shown in Table 5.

Comparison of teacher evaluations of PTC outcomes between the two groups of students

After the PTC was completed, the comprehensive PTC outcomes of the two groups of students were independently evaluated by three attending instructors. Since the data were normally distributed (Shapiro–Wilk test: $p = 0.143$ and $p = 0.421$ for experimental and control groups, respectively) and variances were homogeneous (Levene's test: $p = 0.867$), the independent-samples t-test was selected. The scores for the experimental group and the control group were 21.37 ± 1.23 and 18.93 ± 1.31 , respectively, with $p < 0.001$, indicating a significant difference in PTC performance between the two groups. This showed that live video teaching enhanced the overall dental performance of students, including communicating with the patients in a simulated clinical context, choosing the correct treatments, etc. Specific data are shown in Table 6.

Table 4 Comparison of theoretical comprehensive assessment scores and procedural skills assessment scores between experimental and control groups

Group	Procedural Skills Assessment Scores (points, out of 100; Adjusted Mean \pm SE)	Theoretical Comprehensive Assessment Scores (points, out of 100; Adjusted Mean \pm SE)
Experimental (n=18)	89.49 \pm 0.73	73.10 \pm 1.66
Control (n=18)	79.48 \pm 0.73	69.02 \pm 1.66
p-value	3.39 $\times 10^{-11}$ ($p < 0.001$)	

Table 5 Comparison of Self-Evaluation of learning outcomes between the two groups of students

Survey Content	Full Score	Experimental Group	Control Group	p-value
Stimulating Interest in Learning	5	4.50 \pm 0.62	2.22 \pm 0.73	2.65 $\times 10^{-7}$ ($p < 0.001$)
Improving Logical Thinking Ability	5	4.22 \pm 0.65	2.72 \pm 0.57	2.00 $\times 10^{-6}$ ($p < 0.001$)
Enhancing Procedural Skills	5	4.28 \pm 0.75	2.78 \pm 0.65	1.00 $\times 10^{-5}$ ($p < 0.001$)
Enhancing Doctor–Patient Communication Skills (Simulation)	5	4.11 \pm 0.47	2.83 \pm 0.51	5.44 $\times 10^{-7}$ ($p < 0.001$)
Enhancing Mutual Cooperation Abilities	5	4.50 \pm 0.51	2.56 \pm 0.86	6.02 $\times 10^{-7}$ ($p < 0.001$)
Overall Satisfaction	25	21.61 \pm 1.14	13.05 \pm 1.70	1.36 $\times 10^{-7}$ ($p < 0.001$)

Table 6 Comparison of teacher evaluations of PTC outcomes between the two groups of students

Group	Teacher Evaluation Scores (points)
Experimental (n=18)	21.37 \pm 1.23
Control (n=18)	18.93 \pm 1.31
p-value	2.00 $\times 10^{-6}$ ($p < 0.001$)

EPA assessment scores

Once the PTC was over, students were evaluated by three senior attending physicians who came from the other campus throughout the EPA assessment, and the students also conducted self-evaluations. The combined scores from teacher evaluations and self-evaluations formed the total EPA assessment scores. To assess the effect of the teaching intervention on EPA performance while controlling for baseline differences, a series of one-way ANCOVAs were conducted. For each ANCOVA, the post-PTC EPA score served as the dependent variable, group assignment (experimental vs. control) was the independent variable, and the pre-PTC professional assessment score was the covariate. All necessary statistical assumptions were met for these analyses: scores were normally distributed (Shapiro–Wilk test), error variances were equal across groups (Levene's test), and the homogeneity of regression slopes assumption was satisfied

Table 7 Comparison of EPA assessment scores between the two groups of students

Evaluation Category	Weight	Experimental Group (Adjusted Mean \pm SE)	Control Group (Adjusted Mean \pm SE)	<i>p</i> -value	F(1, 33)	Partial η^2	95% CI for Partial η^2
Self-Evaluation	50%	93.05 \pm 1.26	89.21 \pm 1.26	0.039 ($p < 0.05$)	4.60	0.122	[0.002, 0.344]
Teacher Evaluation	50%	88.79 \pm 0.85	85.86 \pm 0.85	0.021 ($p < 0.05$)	5.87	0.151	[0.003, 0.375]
Overall Evaluation	100%	90.92 \pm 0.91	87.54 \pm 0.91	0.013 ($p < 0.05$)	6.83	0.172	[0.019, 0.385]

(self-evaluation EPA score $p = 0.231$, teacher-evaluation EPA score $p = 0.737$, overall EPA score $p = 0.220$).

The results indicate that after controlling for baseline abilities, the innovative teaching model led to significantly better outcomes across all EPA assessment categories, suggesting that live video teaching is an effective method for developing the PTC performance of students and improving their EPA levels, better preparing them for future internships and clinics. Specific data are shown in Table 7.

Discussion

Our study provides strong evidence that an LTM significantly enhances students' procedural skills and EPA-based assessment scores within a preclinical OMFS course when compared to a traditional teaching model.

Preparing students for the high-stakes clinical environment of OMFS has caused OMFS to undergo significant changes, shifting its focus from treating diseases to enhancing patients' overall well-being and quality of life. This reflects a deeper understanding of effective treatment: addressing clinical issues while improving the patient's overall condition [12]. As the field progresses, physicians and students involved in OMFS must possess a diverse set of skills [13]. These skills not only include theoretical knowledge and clinical abilities, but also the capacity to handle surgical complexities and effectively respond to diverse patient reactions [14]. Addressing this educational challenge requires robust preclinical training that can prepare students for these real-world demands. Consequently, this presents new challenges for our teaching and for students' learning methods, leading to the emergence of EPA evaluation [15].

Through continuous EPA assessments, instructors can comprehensively understand students' procedural skill levels over time, assign different trust levels based on the current stage, and develop personalized training programs. Additionally, EPAs help ensure future patient safety by structuring trainee performance under different trust levels to perform clinical operations [16]. More importantly, EPA evaluations in a preclinical context will allow students to accurately assess their competencies and collaborate with instructors to devise personalized learning plans for future clinical operations [17].

In recent years, EPAs in dentistry have moved from conceptual development to multi-level implementation. At the national guidance level, the ADEA workgroup's

core EPA list and specifications provide a shared language and scope for schools to align curricula and assessments [18]. Programmatic evidence is also accumulating: for instance, curriculum reform that uses EPA-based assessment has successfully supported entrustment decisions before graduation [19]. Additionally, regional pilots have demonstrated the feasibility of tooth-extraction EPAs by using the Delphi method and the EQual rubric on a digital platform [20]. Other studies further indicate broad recognition of the value of EPAs, while highlighting persistent needs in faculty development, institutional support, and mitigation of performance anxiety [21, 22]. Collectively, dentistry now has an emerging consensus baseline and growing evidence for EPAs, yet further optimization is needed in different contexts and rating standards—gaps directly addressed by our EPA-informed PTC design integrating LTM.

Our study's assessment framework is grounded in established international standards. The six competency domains we measured directly align with the core principles of the leading dental education body, the ADEA. This ensures that our evaluation focuses on skills universally considered essential for clinical readiness. Our research contributes to the global implementation of EPAs in two critical ways: (1) we make EPAs practical for a specialty. By focusing on a specific task in OMFS, we show how to apply the broad EPA concept to a targeted, preclinical context, and (2) we provide a validated assessment rubric for the field. By introducing EPAs early in a preclinical course and pairing them with live-video technology, our study provides a replicable model. This approach creates a clear and measurable pathway from foundational training to eventual clinical entrustment.

The TTM within the PTC has several obvious shortcomings when it comes to enhancing the EPA levels of students in OMFS [23]. As a result, firstly, medical education highlights the need to provide repeatable learning methods so that students can practice procedures multiple times and reinforce their skills. Secondly, it is crucial to overcome both fixed viewing angles and limited teaching space by offering multi-angle demonstrations, so that every student can observe the entire procedural demonstration in real time. Finally, it calls for the design of engaging, interactive teaching activities rather than monotonous rote instruction [24].

In the realm of medical education, ongoing research and innovation is needed to improve students'

competencies in EPA, along with collaboration between surgeons, researchers, and engineers [25, 26]. Combining digital teaching technology and the EPA evaluation system on the foundation of traditional teaching methods effectively fulfills our needs to boost conventional teaching approaches [27]. Our results showed that while the experimental group exhibited no significant advantage in theoretical assessment scores compared to the control group, their procedural skills assessment scores were substantially higher, demonstrating that the LTM enormously enhances students' comprehensive procedural abilities and overall performance levels in this pre-clinical study. Considering the EPA scale assessment, the experimental group achieved a higher EPA score versus the scores of the control group, demonstrating that the LTM not only enhanced procedural proficiency but also fosters competencies that are foundational for the principles of independent execution and progressive entrustment. This LTM accelerates students' learning process, alleviates the problem of overemphasizing theory at the expense of practice, and mitigates the disconnection between didactic teaching and procedural skill development. Notably, this enhancement was also reflected in students' self-assessments and teacher evaluations, although these subjective measures need scrupulous interpretation, as discussed in the section on limitations.

The significant improvement in procedural skills can be robustly explained through the lens of cognitive load theory. Surgical procedures such as tooth extraction entail a high intrinsic load due to their inherent complexity. The TTM further burdens students with a high extraneous load, stemming from the "guesswork and visual searching" required to overcome obstructed views [28]. The LTM helps students better integrate theoretical knowledge with practical skills, enabling them to perform better in EPAs and build a stronger foundation to meet the complex demands of their future clinical roles.

Moreover, the successful application and the strong reliability of the EPA scale in this preclinical context indirectly validate its feasibility for this purpose. This clearer scale has the potential to facilitate future entrustment decisions in the entrustment process by providing mentors with objective, verifiable evidence of competency. This ensures that each step in the EPA continuum is underpinned by objective, verifiable evidence of competency.

Limitations

This study has some limitations. First, and most critically, the observed enhancements, such as the self-evaluation and teacher evaluations, may be partially influenced by the Hawthorne effect. The Hawthorne effect suggests that participants' awareness of being studied can alter their behavior, potentially leading to increased motivation in

the experimental group [29]. Similarly, the introduction of novel technology like an LTM can create a 'novelty effect', where the initial excitement and engagement with the new tool have a positive impact on improving performance and satisfaction [30]. Even though our study demonstrates a clear advantage in objective procedural skills, the magnitude should be interpreted with caution. Future research could mitigate these effects by incorporating a longer adaptation period for any new technology.

Second, we acknowledge that instructors could not be blinded to the teaching intervention, which introduces the potential for performance bias. We attempted to mitigate this by rotating all instructors across both groups. In addition, one secondary outcome, the 'teacher evaluation of student PTC outcomes', was intentionally rated by these unblinded instructors to capture their direct pedagogical impressions. This subjective measure is therefore susceptible to bias and should be interpreted with caution, reinforcing the importance of our primary outcomes, which were assessed by blinded external evaluators.

Third, the duration of LTM deployment was relatively brief, which may have resulted in an insufficient variety of data and limited our ability to observe longer-term learning curves. Future studies should extend the tracking time within the EPA framework, in order to assess whether the gains achieved via LTM can be transformed into a sustained boost of entrustment level and practical capabilities in future clinical settings. Researchers can also study whether there is a possibility of applying LTM to clinical internships.

Fourth, while this study established initial evidence for the validity and reliability of our novel OMFS-EPA scale, its psychometric properties require more validation. The current study provided evidence for content validity through expert consensus and preliminary evidence for construct validity via correlations with other performance measures. However, a more robust validation would involve larger and more diverse student cohorts, ideally across multiple institutions. Future research should employ advanced statistical techniques, such as confirmatory factor analysis (CFA). Then, establishing the predictive validity of the scale is a crucial next step.

A significant limitation is the cross-sectional design of this study, which only captured outcomes immediately post-intervention without any long-term follow-up. Consequently, this study cannot confirm the retention of the procedural skills gained over time. More critically, it remains uncertain whether the observed improvements can translate into enhanced performance and higher entrustment decisions in authentic clinical settings. The core purpose of EPA-based preclinical training is to prepare students for real-world responsibilities, and our study lacks the evidence to bridge this crucial gap. To

address this, future research must incorporate a longitudinal follow-up design. Such a study would track the same student cohort into their clinical internships, collecting data on their performance through direct observation by supervisors and formal EPA assessments in the workplace. This would provide the definitive evidence needed to establish the real-world value of the LTM and the predictive validity of our preclinical EPA scale.

Moreover, future research should assess the generalizability of these findings. One avenue is to conduct multicenter studies that replicate this preclinical PTC across diverse educational settings (e.g., different dental schools). Another subsequent line of research could then evaluate the LTM's effectiveness and scalability in actual clinical environments, such as high-resource academic hospitals and lower-resource community clinics. Conversely, these multicenter studies could also provide extensive validation of the OMFS-EPA scale represented in this paper. Second, our sample size was relatively small, which raises concerns about statistical power and generalizability. The sample size reflected the cohort of available students rather than a formal a priori calculation. To address concerns about statistical precision, we conducted a post-hoc power analysis in G*Power (Version 3.1; Heinrich Heine University, Düsseldorf, Germany) for the primary outcome (overall EPA score). Using a two-tailed independent-samples t-test ($\alpha = 0.05$; $n = 18$ per group), the achieved power was approximately 0.75. While this is slightly below the conventional 0.80 benchmark, it still indicates a moderate ability to detect an effect of this magnitude. These findings should nevertheless be interpreted with caution, and larger multicenter studies are warranted to confirm the results. We also report effect sizes and confidence intervals as primary indicators of precision, recognizing that post-hoc power is mainly informative. For future research, the sample range can be expanded by collaborating across colleges or schools to increase the sample size. Additionally, more sophisticated statistical methods can be used to optimize sample selection and data analysis, improving the representativeness and reliability of the results.

Lastly, the high cost of LTM equipment limits scalability in low-income areas. To mitigate this barrier, a multi-pronged strategy can be employed: (1) cost-sharing and money-saving: several institutions can jointly purchase or borrow hardware, or choose low-resolution streaming cameras to save money; (2) public-private collaboration: institutions such as hospitals can cooperate with medical device companies to secure subsidized equipment or donation programs; (3) resource sharing: adopting free, browser-based streaming and analytics tools that only require minimal licensing fees, and building shared resource platforms to promote resource sharing, such as multicenter live-streaming [31].

Conclusion

The integration of an LTM into the OMFS preclinical curriculum offers a significant improvement over traditional teaching methods. This study provides strong evidence that this innovative model enhances students' performance on simulated procedural tasks and improves their overall EPA scores in a preclinical setting. It also boosts students' self-perceived competence and learning engagement. By potentially reducing extraneous cognitive load, this technological intervention helps learners more effectively master complex procedures. Ultimately, this approach builds a stronger foundation for students as they transition to clinical practice, and demonstrates the feasibility of using tailored EPA scales to assess competency development at the preclinical stage.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12909-025-08368-0>.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements

We are grateful to Boyuan 'Blake' Chen for his invaluable assistance with computer maintenance. We also thank Yuxiao Zhu, Shengqiao Li, Sidi Li, Andrea Mirabal, and Gesa Alexes for their dedicated help in maintaining the research equipment. We thank International Science Editing (<http://www.internationalscienceediting.com>) for editing this manuscript. And most importantly, we are thankful to Professor Bao and Dr. Yan for their guidance in this research.

Authors' contributions

Wen Zhe and Lin Yue were responsible for writing the content and performing statistical data analysis. Li Yuqi was responsible for collecting the statistical data. Wei Xiaosong and Zhang Jiafen were responsible for revising and writing parts of the content. Yan Tinglin and Bao Zhexuan provided guidance for writing and publishing the article.

Funding

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

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This paper is approved by the Affiliated Stomatological Hospital of Guangzhou Medical University Ethics Committee, with a registration number of LCYJ2025618007. And this study adhered to the Declaration of Helsinki. Informed consent was obtained from all the participants.

Consent for publication

All participants have consented to the publication.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Comprehensive II, School and Hospital of Stomatology, Guangdong Engineering Research Center of Oral Restoration and

Reconstruction & Guangzhou Key Laboratory of Basic and Applied Research of Oral Regenerative Medicine, Guangzhou Medical University, Guangzhou 510180, China, People's Republic

²Department of Advanced General Dentistry, Yonsei University College of Dentistry, Seoul 03722, Korea, Republic Of

³Department of Periodontology, School and Hospital of Stomatology, Guangdong Engineering Research Center of Oral Restoration and Reconstruction & Guangzhou Key Laboratory of Basic and Applied Research of Oral Regenerative Medicine, Guangzhou Medical University, Guangzhou, Guangdong 510180, China

⁴Department of Dentistry, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong 510150, China

Received: 24 July 2025 / Accepted: 24 November 2025

Published online: 30 December 2025

References

1. Aldweesh AH, Aldhubaiban M, Alqahtani A, et al. Students' perceptions of their dental curriculum and education. *Pak J Med Sci.* 2022;38(6):1441–6.
2. Lee GB, Chiu AM. Assessment and feedback methods in competency-based medical education. *Ann Allergy Asthma Immunol.* 2022;128(3):256–62.
3. Ryan MS, Lomis KD, Deiorio NM, et al. Competency-based medical education in a norm-referenced world: a root cause analysis of challenges to the competency-based paradigm in medical school. *Acad Med.* 2023;98(11):1251–60.
4. Wolcott MD, Quinonez RB, Ramaswamy V, et al. Can we talk about trust? Exploring the relevance of entrustable professional activities in dental education[J]. Volume 84. *JOURNAL OF DENTAL EDUCATION*; 2020. pp. 945–9.
5. Ehlinger C, Fernandez N, Strub M. Entrustable professional activities in dental education: a scoping review. *Br Dent J.* 2023;234(3):171–6.
6. Sethi A, Haq MA, Zaidi SJA, et al. Developing entrustable professional activities for undergraduate operative dentistry clerkship. *BMC Med Educ.* 2024. <https://doi.org/10.1186/s12909-024-06525-5>.
7. Paas F, van Merriënboer JG. Cognitive-load theory : methods to manage working memory load in the learning of complex tasks. *Curr Dir Psychol Sci.* 2020;29(4):394–8.
8. Chen O, Paas F, Sweller J. A cognitive load theory approach to defining and measuring task complexity through element interactivity. *Educ Psychol Rev.* 2023. <https://doi.org/10.1007/s10648-023-09782-w>.
9. Sweller J. The role of evolutionary psychology in our understanding of human cognition: consequences for cognitive load theory and instructional procedures. *Educ Psychol Rev.* 2022;34(4):2229–41.
10. Pinilla S, Lenouvel E, Cantisani A, et al. Working with entrustable professional activities in clinical education in undergraduate medical education: a scoping review. *BMC Med Educ.* 2021. <https://doi.org/10.1186/s12909-021-02608-9>.
11. ADEA Competencies for the New General Dentist[J]. *J Dent Educ.* 2017;81(7):844–.
12. Dorsey ER, Ritzer G. The McDonaldization of medicine. *JAMA Neurol.* 2016;73(1):15–6.
13. Pereira KR, Sinha R. Welcome the “new kid on the block” into the family: artificial intelligence in oral and maxillofacial surgery. *Br J Oral Maxillofac Surg.* 2020;58(1):83–4.
14. Giantini-Larsen AM, Norman S, Pannullo SC. Interns without subinternships. *J Surg Educ.* 2022;79(2):283–5.
15. Schumacher DJ, Turner DA. Entrustable professional activities: reflecting on where we are to define a path for the next decade. *Acad Med.* 2021;96(7S):S1–5.
16. Quinonez RB, Danciu T, Ramaswamy V, et al. Bridging the gap between dental education and clinical practice: the entrustable professional activities model. *J Am Dent Assoc.* 2023;154(8):687–9.
17. Pinilla S, Lenouvel E, Cantisani A, et al. Working with entrustable professional activities in clinical education in undergraduate medical education: a scoping review. *BMC Med Educ.* 2021;21(1):172.
18. Ramaswamy V, Danciu T, Kennedy EN, et al. American Dental Education Association compendium entrustable professional activities workgroup report. *J Dent Educ.* 2024;88(5):639–53.
19. Uyan DC, Tittmore A, Stewart D, et al. Longitudinal assessment using entrustable professional activities: key steps for implementation in dental Education[J]. *JOURNAL OF DENTAL EDUCATION*; 2025.
20. Chen JL, Cheng JH, Wu C, et al. Developing competency-based medical education for dental education in Taiwan: a pilot study of tooth extraction entrustable professional activities. *J Dent Sci.* 2024;19S:128–35.
21. Yousuf AM, Barakaat AA, Tahir K, et al. Evaluating faculty and trainee perceptions on entrustable professional activities in dental postgraduate training: a cross-sectional survey. *BMC Med Educ.* 2025;25:1172. <https://doi.org/10.1186/s12909-025-07366-6>.
22. Cully JL, Schwartz SB, Martini A, Quinonez RB, Harun N, Schumacher DJ, Klein M. Program director attitudes toward acceptance of pediatric dentistry entrustable professional activities (EPAs). *Pediatr Dent.* 2024;46(2):121–34.
23. Burgess A, van Diggele C, Roberts C, et al. Key tips for teaching in the clinical setting. *BMC Med Educ.* 2020;20(Suppl 2):463.
24. Yoneda M, Yamada K, Izumi T, et al. Development of an error-detection examination for conservative dentistry education. *Clin Exp Dent Res.* 2020;6(1):69–74.
25. Bruening DM, Truckenmueller P, Stein C, et al. 360 degrees 3D virtual reality operative video for the training of residents in neurosurgery[J]. *Neurosurg Focus.* 2022;53(2):E4.
26. O'Sullivan S. Models of governance for innovation in medicine and health research. *Eur J Health Law.* 2020;27(3):324–34.
27. Li JO, Liu H, Ting D, et al. Digital technology, tele-medicine and artificial intelligence in ophthalmology: A global perspective[J]. *Prog Retin Eye Res.* 2021;82:100900.
28. Zaed I, Chibbaro S, Ganau M, et al. Simulation and virtual reality in intracranial aneurysms neurosurgical training: a systematic review. *J Neurosurg Sci.* 2022;66(6):494–500.
29. Demetriou C, Hu L, Smith TO, et al. Hawthorne effect on surgical studies. *ANZ J Surg.* 2019;89(12):1567–76.
30. Tsay CH, Kofinas AK, Trivedi SK, et al. Overcoming the novelty effect in online gamified learning systems: an empirical evaluation of student engagement and performance[J]. Volume 36. *JOURNAL OF COMPUTER ASSISTED LEARNING*; 2020. pp. 128–. 2.
31. Yaros J, de Mortier C, Oude EM, et al. Identifying costs in health professions education: a scoping review protocol[J]. *BMJ Open.* 2023;13(10):e74410.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.