applied sciences

MDPI

Article

Motion Tracking-Based Dental Posture Training: Heuristic-Based Assessment and a Comparative Feasibility Study

Jun-Seong Kim [1], Kun-Woo Kim [2], Hyo-Joon Kim [3] and Seong-Yong Moon [1,2,3,*]

[1] Clinical Coordinating Center, College of Dentistry, Chosun University, Gwangju 61452, Republic of Korea; jskim8190@chosun.ac.kr
[2] Dental Biomedical Engineering, College of Dentistry, Chosun University, Gwangju 61452, Republic of Korea; gunwookim3263@gmail.com
[3] Department of Oral and Maxillofacial Surgery, Chosun University Dental Hospital, Gwangju 61452, Republic of Korea; hyojoonkim@chosun.ac.kr
* Correspondence: msygood@chosun.ac.kr

Abstract

Accurate dental working posture is essential for clinical performance and patient safety, yet objective and consistent posture training remains challenging. We present a motion-recognition-based posture assessment approach for dental clinical training that computes posture features from skeletal motion data and applies a rule-based (heuristic) posture assessment module based on predefined clinical criteria to generate posture scores, penalty counts, and warning counts. We conducted a comparative feasibility study with five dental residents (40 sessions total; eight procedure states per participant, 5 min/session). Heuristic-based assessment outputs were compared with specialist video-based ratings using paired nonparametric tests and Bland–Altman analysis. Across pooled sessions (n = 40), the heuristic method yielded lower posture scores than the video-based assessment (89.38 ± 5.21 vs. 97.38 ± 2.99) and higher penalty and warning counts (2.13 ± 1.04 vs. 0.53 ± 0.60; 7.23 ± 3.10 vs. 2.48 ± 1.54). Between-method differences were significant ($p < 0.001$), and Bland–Altman analysis showed a mean bias of 8.00 points (95% limits of agreement: −1.36 to 17.36), indicating limited session-level interchangeability. The proposed approach shows practical feasibility for quantitative posture recording and training support, but larger validation studies and calibration strategies are needed to improve agreement with reference assessments.

Keywords: dental working posture; motion tracking; heuristic-based posture assessment; posture training; virtual reality training

Check for updates

Academic Editor: Chaman Sabharwal

Received: 9 February 2026
Revised: 26 February 2026
Accepted: 30 June 2026
Published: 2 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Background of the Study

Advances in technology are changing job requirements across a wide range of occupations, and adapting to these changes requires training a workforce equipped with the necessary competencies [1–3]. Vocational education supports this need by promoting both theoretical understanding and practical performance skills that can be applied directly in the workplace [1–3].

However, practice-oriented training may be limited in sustainability and efficiency due to cost, space, equipment availability, and safety constraints [4,5]. Establishing and maintaining training environments that closely resemble real-world settings often requires substantial resources, and it can be difficult to rapidly incorporate technology-driven

changes into existing programs. When hands-on training is conducted in real workplaces, additional issues may arise, including safety risks for novices and reduced learning opportunities due to limited access to tools and equipment. To mitigate these challenges, virtual reality (VR) has been increasingly integrated into vocational education because it enables standardized, repeatable practice in a safe environment [4–8].

Importantly, the assessment target determines the methodological requirements of VR training systems. While many VR programs emphasize procedural execution and outcome-oriented performance metrics, posture training requires continuous monitoring of whole-body kinematics and conversion of posture features into interpretable, criteria-based feedback. This posture-oriented assessment introduces distinct challenges, such as motion-tracking accuracy and latency, calibration and coordinate alignment, and robustness to transitional movements (e.g., repositioning or tool switching) that can confound posture measurements [9–11].

1.2. Related Work on VR-Based Vocational Education

Recent research on VR-based vocational education has expanded toward competency-based simulations that allow learners to perform real work procedures step by step in a virtual space, aiming to reduce constraints related to training cost, safety, and repeatability [9,10,12]. In skilled technical domains, VR-based welding training has been reported to reduce safety risks for novice learners while providing real-time feedback on key parameters (e.g., torch angle, travel speed, and working distance), thereby improving task performance and self-confidence [11,13,14]. VR training platforms have also been developed for high-risk and high-cost tasks, such as aircraft maintenance and industrial equipment maintenance, with studies reporting improvements in procedural proficiency and operational efficiency [15–20]. In addition, VR applications have broadened to include assembly and disassembly training, including platforms that support human–robot disassembly training and synthetic data generation [21,22].

VR has also been actively adopted in industrial safety education in sectors such as construction and logistics. Immersive VR frameworks that incorporate interactive learning elements (e.g., immediate feedback, object manipulation/assembly, and knowledge tests) and scenario-based approaches for tasks such as forklift operations and human–robot collaboration have been proposed [23–25]. Beyond providing workplace-like experiences, VR systems can generate assessable logs and metrics (e.g., time, errors, and events) to support training analysis and objective evaluation [10,23].

In medical education, VR has likewise been increasingly used to enable safe, repeatable practice without exposing patients to avoidable risks [26,27]. Early and widely referenced applications include commercial VR platforms and case studies for anatomy education, multi-user collaborative training, and oral/maxillofacial surgery contexts [28–31], and more recent work has expanded VR-based procedural skills training with evidence from controlled studies, including randomized trials and longer-term follow-up reports [27,32]. In surgical and interventional training, immersive VR has been reported to improve technical performance as well as learning immersion, satisfaction, and self-efficacy [33]. Portable and affordable VR surgical simulators have also been shown to improve module-specific performance and overall efficiency [34]. VR simulations have further been explored for complex procedures, including neurosurgical ventriculostomy simulation and maxillofacial procedures such as Le Fort I orthognathic surgery [35,36]. Many VR-based medical training systems evaluate skills using quantitative indicators generated during task execution (e.g., time, errors, procedural steps, and accuracy) [26,37–43], with continued development across broader procedural domains such as emergency/team-based education, airway management, ultrasound-guided procedures, and vascular interventions [44–52]. VR-based

training has also been applied across multiple professions and educational stages, including nursing, endoscopy, and radiology/radiography education, and VR simulators have been evaluated for usability in emergency procedures such as umbilical venous catheter (UVC) placement [53–59]. In dentistry, VR (and haptic) simulators and simulator-derived metrics have been explored as potential indicators related to outcomes in conventional model-based training, and systematic reviews have discussed the educational role of virtual haptics in surgical simulation [60,61]. Collectively, these studies suggest that VR is becoming a core educational medium supporting standardized repetitive practice and quantitative assessment in healthcare education [26], with growing evidence regarding feasibility and assessment reliability in OSCE-oriented procedures and clinical examination training [62–64].

1.3. Research Gap and Objectives

Despite the rapid growth of VR-based medical education, VR programs that explicitly target dental clinical working posture training and assessment remain relatively underexplored. This gap is important because dental working posture is not only an ergonomic issue but also a clinically relevant factor associated with operator fatigue, musculoskeletal load, and treatment performance [65,66]. Given the high prevalence of work-related musculoskeletal disorders among dental professionals, posture training is widely regarded as an essential component of dental education, particularly because repetitive motions, prolonged static postures, and non-ergonomic working positions are common in dental practice [67,68]. However, repeated posture training and consistent evaluation are often difficult to provide in real clinical or practice settings due to constraints in space, equipment availability, time, and patient-related ethical and safety considerations.

Most VR/simulation-based healthcare training systems primarily assess procedural completion and outcome-oriented metrics (e.g., time, errors, and accuracy). In contrast, posture-oriented assessment requires continuous extraction of posture features from motion-tracking data and mapping these features to predefined clinical posture criteria to produce transparent and interpretable feedback (e.g., scores and warnings). Because transitional movements (e.g., repositioning or tool switching) can introduce short-lived deviations, posture scoring may become overly sensitive or unstable unless transitions are explicitly handled. Unlike prior VR dental simulators that mainly report task performance outcomes, our system converts motion-tracking kinematics into criterion-linked scores and warnings and evaluates these outputs against specialist-based ratings. Therefore, there is a need for dental VR training systems that operationalize criterion-linked posture assessment and examine how their outputs behave relative to specialist-based ratings in a comparative evaluation setting.

In this paper, we propose a VR-based solution for dental clinical posture training using motion-tracking devices. Our approach aims to: (i) construct a VR clinical environment that supports dental procedure practice; (ii) visualize trainees' skeleton-based motion data in VR and compute posture-related features; and (iii) develop a rule-based (heuristic) posture assessment module grounded in clinical posture criteria to provide real-time, interpretable feedback. We adopt a heuristic approach because it directly links measured features to predefined criteria, enabling transparent educational feedback without requiring large labeled datasets. At the same time, we acknowledge its trade-offs, including sensitivity to threshold settings, limited personalization across trainees, and potential false warnings during transitional movements. Finally, we conduct a comparative feasibility study with dental residents to evaluate the educational applicability and practical feasibility of the proposed system by comparing its outputs with specialist-based ratings.

2. Materials and Methods

This section describes the development workflow and implementation methods of the proposed VR-based dental posture training system and presents a human-subject comparative feasibility study to evaluate the 1 module. The study consists of two components: (i) system development, including VR environment construction, motion visualization, and heuristic assessment design, and (ii) a feasibility study with dental residents, in which heuristic assessment outputs were compared with specialist video-based ratings. As summarized in Figure 1, the pipeline comprises VR scenario construction, acquisition of trainees' skeletal motion data using a motion-tracking device, posture visualization and feature computation within VR (e.g., distances and joint/segment angles), and heuristic-based posture scoring by comparing computed metrics with predefined clinical posture criteria. The feasibility-study protocol (participants, procedures, and comparative evaluation) is described in Section 2.3.

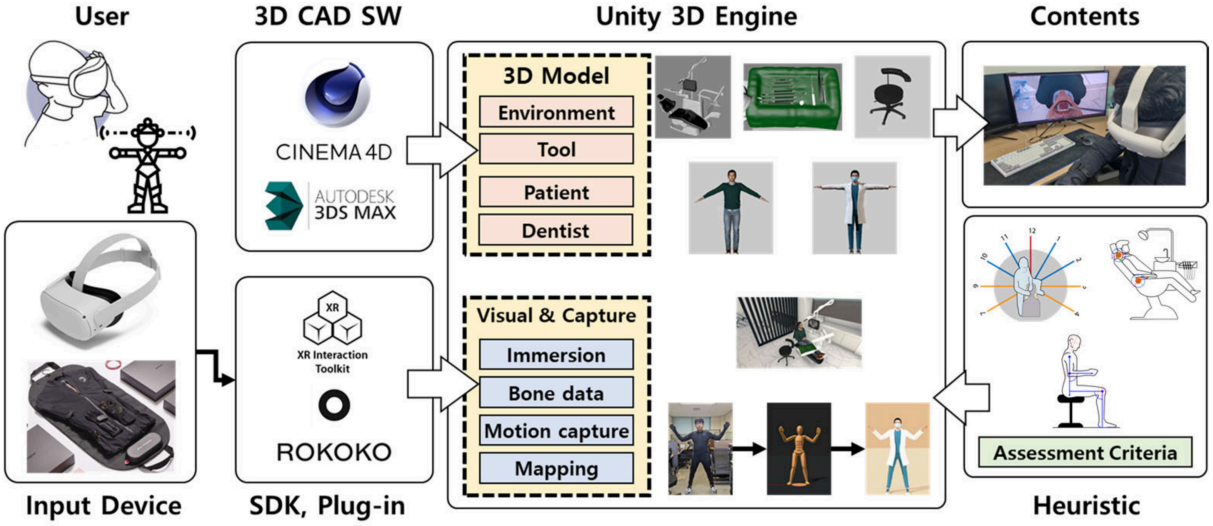


Figure 1. Overall architecture of the proposed VR-based dental posture training solution.

2.1. Virtual Environment Development and Motion Visualization

This subsection describes the system-development component of the proposed solution, including virtual environment construction, human/avatar modeling, and motion-tracking-based posture visualization. The human-subject comparative feasibility study is described in Section 2.3.

To enhance immersion and learning effectiveness, we constructed a virtual clinical environment as the training space. We first conducted an on-site survey of the real clinical setting to identify the layout, frequently used equipment, and procedure-related instruments. Based on the survey, 3D assets were created using 3Ds Max 2025.3 and Cinema4D R23.110 and were grouped into three categories: (i) environment models, (ii) equipment and instrument models, and (iii) human models.

The environment models defined the spatial boundaries and ambience of the virtual clinic (e.g., walls, patient waiting area, and decorative elements). To reduce unintended object interference during interaction, collision handling was implemented in Unity by configuring colliders and controlling inter-layer collisions via the Physics Layer Collision Matrix. Colliders were assigned according to object roles (static structures vs. interactive objects), and collisions between irrelevant layers were disabled to minimize unnecessary contacts while preserving essential interactions. Figure 2 shows the constructed virtual clinical environment.



Figure 2. The constructed virtual clinical environment.

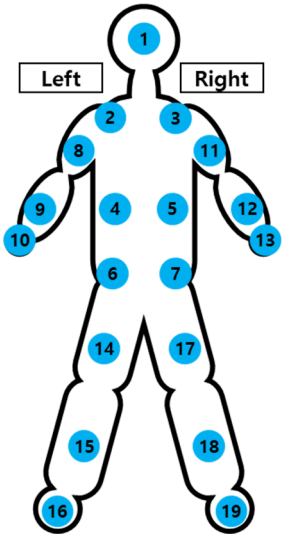
For equipment and instrument models, we created assets for items commonly used in clinical practice (e.g., patient chair, dentist chair, and procedure-related instruments). The models were designed to closely match real-world dimensions based on specifications and direct measurements where applicable. Mesh optimization and texture refinement were performed to maintain visual fidelity while enabling real-time rendering. In Unity, interaction components (e.g., Colliders and Rigidbodies) were assigned to support selection and manipulation with physically plausible responses. These assets were integrated into the VR environment to support object interaction and procedure-state reproduction in the training pipeline.

The human models consisted of a patient model and a dentist model. To synchronize the patient model with the patient-chair model, we defined a skeletal structure for the head, neck, and torso so that the upper body moves naturally with chair rotation and tilt. To support posture adjustments across maxillary and mandibular procedures, we assigned a left–right rotational range of motion to the neck joint. The dentist model was rigged with a hierarchical joint tree (e.g., hands, arms, and upper body) to support motion retargeting from an HMD and motion-tracking devices. This configuration enables natural reproduction of trainees' posture changes and facilitates posture visualization within the virtual environment.

The dentist avatar rig was integrated with full-body motion data captured using the SmartSuit Pro II (Rokoko Electronics ApS, Copenhagen, Denmark). The streamed data provided joint pose information (position and rotation) for major trunk and limb joints; in this study, pose data for 19 joints were used. Joint/segment labels followed the Rokoko naming convention (e.g., Arm, ForeArm, Thigh, and Shin), and a correspondence map was defined between Rokoko skeletal nodes and Unity avatar rig nodes. The received joint transforms were then retargeted to the avatar based on this mapping. Figure 3 shows an example of the SmartSuit skeletal data, and Figure 4 shows the motion-synchronized virtual dentist model.

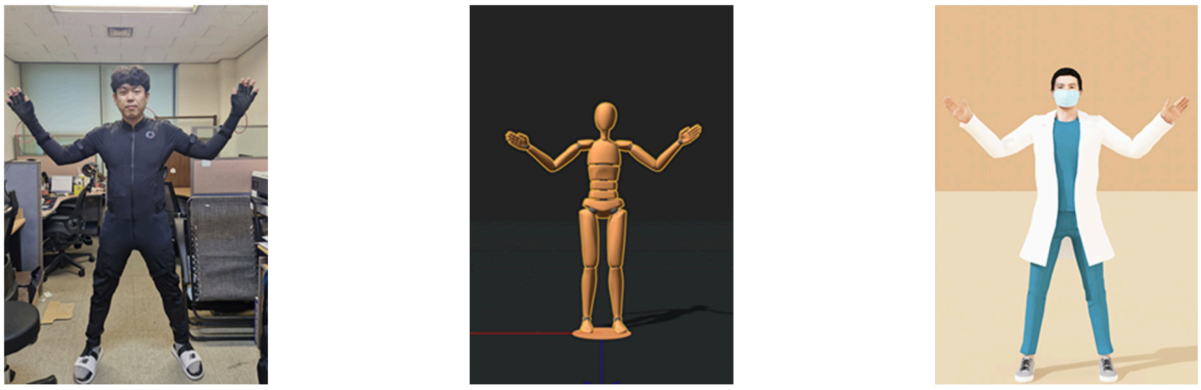
The SmartSuit Pro II measures sensor-based full-body joint poses, and its joint hierarchy and coordinate system are predefined by the device. Before applying incoming poses, the streamed transforms were converted to match the avatar rig format and reoriented to the Unity world coordinate system. At the beginning of each session, calibration was initialized by holding a standardized standing pose (idle/attention pose) or a T-pose for approximately 3–5 s, following the Rokoko calibration procedure. Using this initial pose, joint-specific rotational offsets were computed and applied during retargeting so that the SmartSuit joint rotations align with the avatar rig's reference pose. To mitigate sensor-related jitter and short-lived fluctuations that may cause spurious threshold crossings, we stabilized the computed posture metrics (e.g., angles, distances, and clock-position

azimuths) prior to threshold comparison in the heuristic assessment. Specifically, we applied a median filter (≈0.33 s window) to suppress impulsive spikes, followed by a first-order low-pass filter with a fixed cutoff frequency of 4 Hz to reduce high-frequency jitter. The stabilized metric stream was subsequently used for posture-feature computation and heuristic assessment, as described in Section 2.2. In the current implementation, transitional movements (e.g., repositioning or tool switching) were not explicitly segmented or excluded from scoring. Accordingly, the metric smoothing described above was applied prior to threshold comparison to reduce short-lived transition-induced fluctuations.



Index	Bone Name
1	Head
2, 3	Shoulder
4, 5	Spine
6, 7	Hip
8, 11	Arm
9, 12	ForeArm
10, 13	Hand
14, 17	Thigh
15, 18	Shin
16, 19	Foot

Figure 3. Skeletal data provided by the SmartSuit Pro II.



(**a**) (**b**) (**c**)

Figure 4. SmartSuit Pro II-driven avatar synchronization: (**a**) real-world pose; (**b**) Rokoko Studio (v 2.5.1.1); (**c**) synchronized virtual model.

2.2. Heuristic Design for Posture Training

Posture assessment in dental clinical training requires simultaneous consideration of multiple factors, including working distance, head and upper-limb joint angles, operator clock position, and patient chair positioning, which makes consistent and immediate evaluation challenging. To address this, we designed a heuristic assessment that converts guideline-based posture criteria into quantitative metrics (distance, angle, and orientation) using skeletal data from a motion-tracking device and reference points from the virtual patient model. Each criterion is evaluated by checking whether the computed metric falls within a predefined threshold interval. Figure 5 shows the posture concepts used for evaluation (basic posture, clock position, and chair position), and Table 1 summarizes the corresponding rule set.

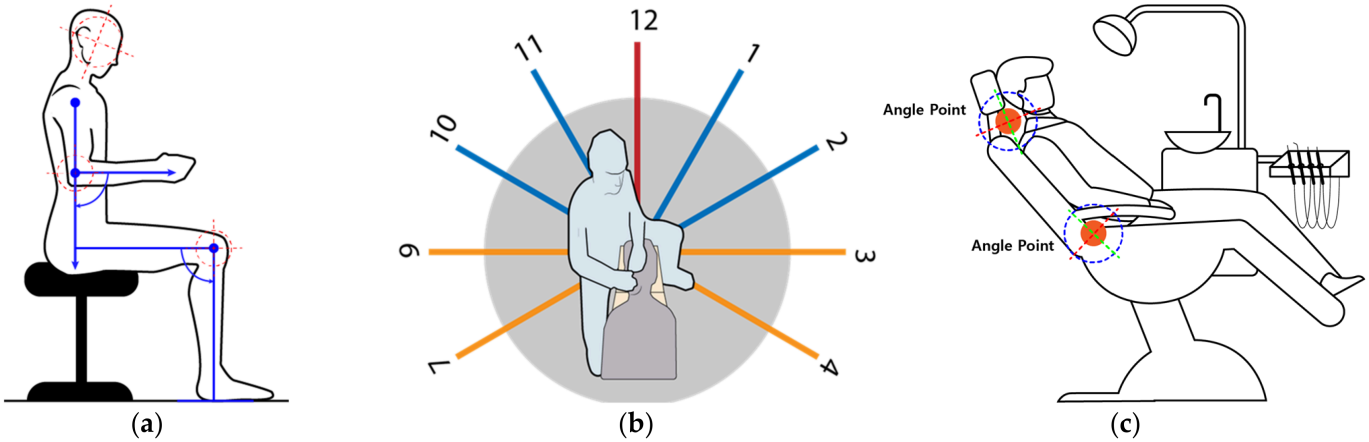


Figure 5. Clinical posture assessment criteria: (**a**) basic posture; (**b**) clock position; (**c**) chair position.

Table 1. Posture evaluation criteria and heuristic threshold ranges.

Type		Item	Baseline of Assessment	Reference Point *5	Ranges
Basic	Distance	W. * point–O. ** eye	250~300 mm	$p_{head(h)}$, $pp_{headrest(hr)}$	$250 < d_{ph-pphr} < 300$ *6
	Angle	W point–O eye	20~30°	$p_{shoulder\ center(sc)}$, p_h, pp_{hr}	$20 < \theta_{mp} < 30$ *7
		O wrist	Wrist kept level	$p_{elbow(e)}$, $p_{wrist(w)}$, $p_{hand(ha)}$	$0 < \theta_w < 50$
		O elbow	Kept close to the torso	$p_{spine\ center(spc)}$, p_{sc}, p_e	$0 < \theta_{sc} < 50$
		Thigh level	Upper–lower leg perpendicular	p_{sc}, $p_{spine\ center(spc)}$, p_e	$80 < \theta_{spc} < 100$
		Hip abduction	30°	$p_{hip(h)}$, $p_{knee(bone-bone)}$, $p_{foot(f)}$	$25 < \theta_{knee} < 35$
	Usage	Left	Mirror required (Y [0]/N [0])	*not reference*	0 or 1
		Right	Operative instrument required (Y/N)	*not reference*	0 or 1
Clock	General operator position		9–12 o'clock	$p_{hip\ cente(hc)}$, pp_{hr}, wp_y *8	$-30 < \theta_{hrclock} < 30$
	Restorative dentistry		11–12 o'clock		$-30 < \theta_{hrclock} < 0$
	M. & M. A.***; M. R. M. **** (direct vision)		7 o'clock		$-150 < \theta_{hrclock} < -120$
	M. & M. A. ***; M. R. M. **** (occlusal surface)		9 o'clock		$-90 < \theta_{hrclock} < -60$
	M anterior lingual surfaces		12 o'clock		$0 < \theta_{hrclock} < 30$
Chair	Chair backrest	Maxillary	Horizontal relative to the floor	pp_h, pp_{hc}, wp_y	$80 < \theta_{backrest} < 95$
		Mandibular	45° relative to the floor		$40 < \theta_{backrest} < 50$
	Headrest	Maxillary	Horizontal relative to the floor	pp_h, $pp_{neck(n)}$, wp_y	$80 < \theta_{headrest} < 95$
		Mandibular	45° relative to the floor		$40 < \theta_{headrest} < 50$
	Patient Head	Patient head Right	Toward the patient's line of sight	pp_{hr}, pp_h, wp_y	$5 < \theta_{pph} < 30$
		Patient head Left			$-5 < \theta_{pph} < -30$
		Patient head Up	Adjust to the dentist's chair height	p_{hc}, p_f, pp_h, wp_y	$(d_{pph-wpy} - 200) < d_{pf-phc} < (d_{pph-wpy} + 100)$
		Patient head Down			

* W. = Working; ** O. = Operator; *** M. & M. A. = Maxillary & mandibular anterior labial surfaces; **** M. R. M. = Mandibular right molars; *5 operator point = p, patient point = pp, unity point = wp, except for the center point, left/right labels are omitted.; *6 See Equation (2) for details; *7 See Equation (4) for details.; *8 See Equation (6) for details.

To compute the posture metrics in Table 1, we selected the minimal set of tracked joints/points required for geometric computation while preserving interpretability. Distance-based criteria were computed using two points, whereas angle-based criteria were computed using three points to form two vectors and estimated using a dot-product formulation. For criteria requiring a fixed reference axis or plane (e.g., chair/headrest tilt or patient-head rotation), we used stable reference frames from the Unity world coordinate system and/or quasi-static reference points on the virtual patient model. Because some composite landmarks required for computation are not provided by the SmartSuit output (e.g., shoulder center and hip center), we deterministically derived these points from available joints (e.g., midpoints between left and right shoulder/hip joints) and used them as reference points.

In this study, the posture criteria were constructed using a two-step process. First, posture-assessment items and baseline recommendations were identified from published dental clinical/educational posture guidelines [69–71]. Second, these recommendations were operationalized into quantitative threshold intervals for VR-based motion analysis. For criteria with clearly reported ranges, the published ranges were adopted directly. For criteria described qualitatively, ambiguously, or as single target values (e.g., clock-position categories or hip abduction), threshold intervals were defined or refined through expert consensus to support stable quantitative assessment. Specifically, two dental experts (co-authors) discussed these items and set tolerance margins by examining how computed angles and distances fluctuate with small natural movements under a controlled baseline setup (a simplified VR environment with a minimal avatar and only the kinematic computation enabled). This refinement accounted for measurement uncertainty and inter-individual variability while remaining consistent with the guideline intent [69–71].

These criteria are meaningful in both training and clinical contexts because they relate to operator biomechanical load, patient head stability, and maintenance of an adequate visual field. Because the criteria are expressed in measurable forms (angle, distance, orientation, and position), they can be directly converted into a motion-tracking-based quantitative assessment algorithm. Accordingly, Table 1 summarizes the rule-design basis (guideline-derived and expert-consensus–operationalized thresholds) rather than participant-derived tuning. Participant characteristics and the feasibility-study protocol are described separately in Section 2.3.

The heuristic assessment computes distances and angles from 3D joint coordinates acquired by the motion-tracking device and determines posture appropriateness by comparing these measurements against predefined threshold intervals. All joint coordinates are represented in the Unity world coordinate system and are used for distance/angle computation and threshold-based decision making. We represent each joint coordinate as:

$$\boldsymbol{p}_i = (\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{z}_i), \text{ where } i \text{ denotes the joint index.} \quad (1)$$

Based on the virtual-space coordinate representation, Euclidean distance is used to compute inter-point distances, and the cosine angle based on the vector dot product is used to compute joint angles. For distance computation, two points are used, and the distance is calculated using 2 of an equation:

$$d(\boldsymbol{p}_i, \boldsymbol{p}_j) = \|\boldsymbol{p}_i - \boldsymbol{p}_j\| = \sqrt{(\boldsymbol{x}_i - \boldsymbol{x}_j)^2 + \left(\boldsymbol{y}_i - \boldsymbol{y}_j\right)^2 + (\boldsymbol{z}_i - \boldsymbol{z}_j)^2}, \quad (2)$$

for example, to compute the distance between the dentist model's head center and the patient's oral-cavity center, Equation (2) is evaluated by setting $\boldsymbol{p}_i$ to the dentist model head-

center coordinate and p_j to the patient oral-cavity-center coordinate, and then calculating the Euclidean distance between the two points.

Before computing joint angles, direction vectors are defined from pairs of reference points in the virtual space. Given two points p_i and p_j, the relative vector is expressed as Equation (3):

$$\mathbf{v}_{ij} = p_i - p_j. \tag{3}$$

Angles were computed using the dot-product-based cosine formulation. For two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, the angle is θ obtained as (4):

$$\theta = \cos^{-1}\left(\frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\|\|\mathbf{v}_2\|}\right), \tag{4}$$

for example, to compute the elbow flexion angle, we use the 3D joint coordinates of the shoulder (p_{sh}), elbow (p_{el}), and wrist (p_{wr}). Taking the elbow as the reference point, we define two relative vectors: $\mathbf{v}_{upper} = p_{sh} - p_{el}$ from the elbow to the shoulder and $\mathbf{v}_{lower} = p_{wr} - p_{el}$ from the elbow to the wrist. The elbow flexion angle θ_{elbow} is then obtained by substituting these vectors into Equation (4), i.e., setting $\mathbf{v}_1 = \mathbf{v}_{upper}$ and $\mathbf{v}_2 = \mathbf{v}_{lower}$. The resulting angle is compared against a predefined threshold interval and used to determine whether the elbow posture is acceptable at each frame. Patient posture (e.g., chair backrest angle, headrest angle, and head rotation) was quantified using the same principle as operator posture. For each component, a direction vector (or the occlusal-plane normal vector) was defined, and the posture state was expressed by computing the angle with respect to reference axes (e.g., the Unity world coordinate axes).

To compute the operator's relative position (clock position), we used a headrest-based reference point of the patient p_{hr}, and an operator reference point (e.g., hip) p_{hip}. Using Equation (3), the operator–patient vector is computed as $\mathbf{v}_{clock} = p_{hip} - p_{hr}$. Because the clock position is defined by the direction on the horizontal plane, we projected $\mathbf{v}_{clock}$ onto the horizontal plane by removing its component along the plane normal. Using the unit normal vector of the horizontal plane in the Unity world coordinate system, $\mathbf{v}_{unity_Y} = (0,1,0)$, the projected vector is computed as Equation (5):

$$\mathbf{v}'_{clock} = \mathbf{v}_{clock} - (\mathbf{v}_{clock} \cdot \mathbf{v}_{unity_Y})\mathbf{v}_{unity_Y}, \tag{5}$$

by applying the projection in Equation (5), the vertical component of $\mathbf{v}_{clock}$ is removed, yielding the horizontal-plane projected vector $\mathbf{v}'_{clock} = (x_{clock}, 0, z_{clock})$. The azimuth angle ϕ (0–360°) is then calculated from the projected vector $\mathbf{v}'_{clock}$ and categorized into predefined angular intervals to determine the clock position classes, as expressed in Equation (6):

$$\phi = \mathbf{atan2}(x_{clock}, z_{clock}) \times \frac{180}{\pi}, \phi \leftarrow (\phi + 360) \bmod 360. \tag{6}$$

Using the above equations, all posture-related metrics (e.g., distances, joint angles, and clock-position azimuths) are computed from the tracked 3D joint coordinates and reference points in the virtual environment. These computed values are subsequently compared with predefined threshold intervals to determine whether each posture criterion is satisfied. Figure 6 shows a simple example of how representative distances and angles are computed in our framework.

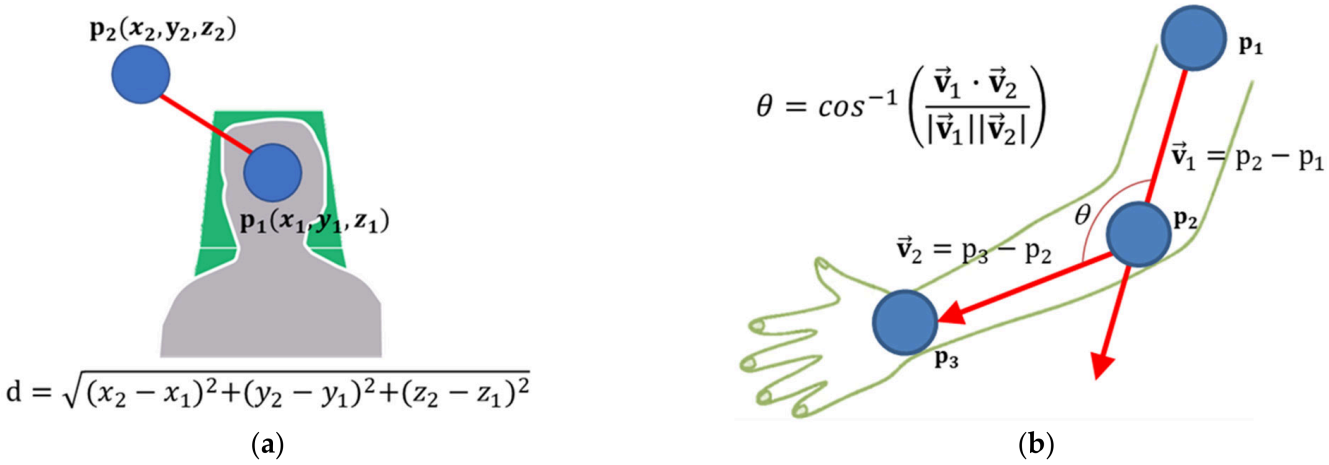


Figure 6. Example of distance and angle calculations used in the heuristic method: (**a**) distance; (**b**) angle.

2.3. Comparative Feasibility Study

Following the system development and posture-assessment design described in Sections 2.1 and 2.2, we conducted a comparative feasibility study with dental residents to evaluate the proposed heuristic posture assessment. We compared conventional video-based posture assessment with the proposed heuristic assessment to examine: (i) the extent to which posture judgments produced by the system align with specialist ratings and (ii) whether real-time quantitative assessment can be practically used in an educational setting. Accordingly, this study focused on feasibility and immediate assessment performance rather than longitudinal posture improvement over repeated training sessions.

Five second-year dental residents with prior clinical training experience participated (n = 5). Participants were recruited by direct invitation, and participation was voluntary. Written informed consent was obtained from all participants. No personally identifiable information was collected; only anonymized assessment outcomes were recorded. The threshold intervals used in the heuristic method were predefined from the clinical posture criteria described in Section 2.2 and were not tuned using participant characteristics. Because this was a feasibility study, the sample was a small convenience sample intended for practicality assessment and between-method comparison rather than population-level generalization.

Each trial lasted 5 min to standardize evaluation conditions and to reduce posture distortion due to cumulative fatigue. Before each session, participants received a short procedure-state description specifying the assigned task. Each participant performed eight procedure states (40 sessions total), defined as four procedure types (scaling, root planing, crown preparation, and extraction) combined with two arch conditions (maxillary and mandibular). Participants performed the required posture for each state using the designated representative target location as the reference. Outcome measures (score, warning count, and penalty count) were recorded for each participant and state to examine between-participant variability descriptively.

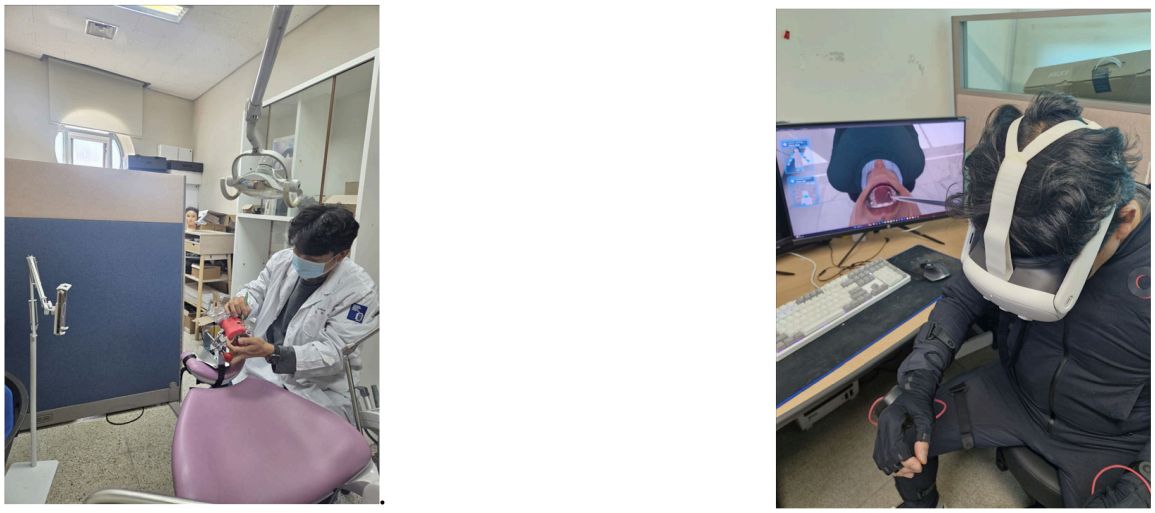
For the conventional video-based assessment, participants were recorded using two smartphone cameras placed at frontal and lateral viewpoints to reduce occlusions. Two dental specialists (co-authors) jointly reviewed each session and assigned a single consensus rating based on discussion using posture guidelines commonly applied in clinical education (basic posture, clock position, and chair position). To enable direct comparison between methods, the same evaluation categories were used in both the specialist ratings and the heuristic method. The specialists were blinded to participant identity and session metadata during video review, except that all participants were second-year residents by study design. Because the final video-based outcome was produced as a consensus score

rather than two independent ratings, inter-rater reliability statistics (e.g., ICC/κ) were not computed in this study.

For the heuristic assessment, the system computed posture metrics in real time using (i) 3D skeletal joint coordinates captured by the motion-tracking device and (ii) reference points from the virtual patient model. Each criterion was evaluated by comparing computed measurements with the predefined threshold intervals (Section 2.2).

The same scoring logic was applied to both methods to ensure direct comparability. Criterion violations were accumulated as warnings, one penalty was applied for every three warnings, and the final posture score was calculated from a maximum of 100 points by subtracting 5 points per applied penalty.

Because of discomfort associated with device wearing and differences in recording and assessment procedures, the video-based and heuristic assessments were not performed simultaneously within the same session. Instead, each method was conducted separately under the same procedure-state conditions. Between-method differences in final posture score, warning count, and penalty count were analyzed to examine the extent to which the proposed heuristic assessment could reproduce the outcomes of the conventional video-based assessment under standardized conditions. This study was designed to evaluate between-method agreement and practical feasibility in a controlled VR training environment. Figure 7 shows posture assessment using each method.



(**a**) (**b**)

Figure 7. Exec. ion of each assessment method: (**a**) Video-based; (**b**) Heuristic-based.

3. Results

3.1. Comparative Outcomes Between Methods

Table 2 shows the comparative outcomes of the video-based and heuristic-based assessments. For the comparative analysis, the session was used as the unit of analysis (40 sessions per method), and the compared outcomes included final posture score, warning count, and penalty count.

Across the 40-session analysis, the heuristic-based assessment consistently produced lower scores than the video-based assessment (heuristic: 89.38 ± 5.21, 95% CI 87.71–91.04, range 80–100; video: 97.38 ± 2.99, 95% CI 96.42–98.33, range 90–100). The session-level score difference (video-heuristic) averaged 8.00 ± 4.78 points (95% CI 6.47–9.53, range 0–20), with the largest gap observed for extraction (#28) in P4 (video 100 vs. heuristic 80; 20-point difference).

Table 2. Comparison results of posture assessment between the two methods.

Procedure Type	Tooth No.	Method	P1			P2			P3			P4			P5		
			S. *	P. **	W. ***	S.	P.	W.	S.	P.	W.	S.	P.	W.	S.	P.	W.
Scaling	#11 *4	Heuristic	85	3	11	90	2	6	90	2	7	95	1	3	95	1	4
		Video	95	1	4	95	1	3	95	1	3	100	0	1	95	1	3
	#36	Heuristic	90	2	7	90	2	7	90	2	6	85	3	10	90	2	7
		Video	95	1	3	100	0	2	100	0	2	95	1	3	95	1	5
Root planing	#27	Heuristic	85	3	11	80	4	13	100	0	2	90	2	7	85	3	10
		Video	95	1	3	95	1	4	100	0	0	100	0	2	95	1	4
	#15	Heuristic	80	4	13	85	3	10	95	1	4	95	1	4	95	1	3
		Video	95	1	3	100	0	1	100	0	1	100	0	1	95	1	3
Crown prep	#25	Heuristic	85	3	10	90	2	6	85	3	10	100	0	2	90	2	6
		Video	90	2	7	100	0	0	95	1	3	100	0	1	100	0	2
	#45	Heuristic	85	3	9	90	2	6	85	3	9	90	2	6	100	0	2
		Video	95	1	3	100	0	1	95	1	4	100	0	1	100	0	2
Extraction	#47	Heuristic	85	3	11	90	2	7	85	3	9	90	2	7	95	1	3
		Video	95	1	4	100	0	2	100	0	2	100	0	2	95	1	3
	#28	Heuristic	85	3	10	90	2	8	90	2	7	80	4	12	95	1	4
		Video	90	2	6	100	0	1	100	0	0	100	0	2	100	0	2
Mean ± SD		Heuristic	85.00 ± 2.67	3.00 ± 0.53	10.25 ± 1.75	88.13 ± 3.72	2.38 ± 0.74	7.88 ± 2.47	90.00 ± 5.35	2.00 ± 1.07	6.75 ± 2.71	90.63 ± 6.23	1.88 ± 1.25	6.38 ± 3.42	93.13 ± 4.58	1.38 ± 0.92	4.88 ± 2.64
		Video	93.75 ± 2.31	1.25 ± 0.46	4.13 ± 1.55	98.75 ± 2.31	0.25 ± 0.46	1.75 ± 1.28	98.13 ± 2.59	0.38 ± 0.52	1.88 ± 1.46	99.38 ± 1.77	0.13 ± 0.35	1.63 ± 0.74	96.88 ± 2.59	0.63 ± 0.52	3.00 ± 1.07
95% CI [L–U]		Heuristic	82.77–87.23	2.55–3.45	8.78–11.72	85.01–91.24	1.75–3.00	5.81–9.94	85.53–94.47	1.11–2.89	4.48–9.02	85.41–95.84	0.83–2.92	3.52–9.23	89.30–96.95	0.61–2.14	2.67–7.08
		Video	91.81–95.69	0.86–1.64	2.83–5.42	96.81–100.69	−0.14–0.64	0.68–2.82	95.96–100.29	−0.06–0.81	0.66–3.09	97.90–100.85	−0.17–0.42	1.00–2.25	94.71–99.04	0.19–1.06	2.11–3.89

* S. = Score; ** P. = Number of Penalty; *** W. = Number of Warning; *4 #number = indicates the tooth number.

Penalty counts were also higher in the heuristic-based assessment (2.13 ± 1.04, 95% CI 1.79–2.46, range 0–4) than in the video-based assessment (0.53 ± 0.60, 95% CI 0.33–0.72, range 0–2). In the video-based assessment, penalty counts were distributed as 0 in 52.5% (21/40) sessions, 1 in 42.5% (17/40), and 2 in 5.0% (2/40). In contrast, the heuristic-based assessment showed a broader distribution: 0 (7.5%, 3/40), 1 (17.5%, 7/40), 2 (37.5%, 15/40), 3 (30.0%, 12/40), and 4 (7.5%, 3/40). Warning counts were likewise higher with the heuristic-based assessment (7.23 ± 3.10, 95% CI 6.23–8.22, range 2–13) than with the video-based assessment (2.48 ± 1.54, 95% CI 1.98–2.97, range 0–7).

By procedure type, video-based mean scores remained high (scaling 96.5, root planing 97.5, crown preparation 97.5, extraction 98.0), whereas heuristic-based mean scores were lower (90.0, 89.0, 90.0, and 88.5, respectively). Consistent trends were observed for penalty and warning counts (mean penalties: heuristic 2.0–2.3 vs. video 0.4–0.7; mean warnings: heuristic 6.3–8.2 vs. video 2.0–3.1), indicating that the heuristic-based method yielded stricter assessments across tasks.

From a training perspective, this stricter, criterion-driven behavior may be beneficial for formative coaching by prompting early correction of posture deviations in real time. At the same time, the increased warning frequency suggests that brief deviations (e.g., transient movements or borderline fluctuations) may be penalized more often; therefore, the outputs should be interpreted primarily as training feedback rather than as high-stakes summative judgment.

3.2. Between-Method Differences and Agreement Analysis

In this section, we evaluated between-method differences and agreement between the video-based and heuristic-based assessments using paired inferential tests, nonparametric paired tests for discrete outcomes, and Bland–Altman analysis with proportional-bias linear regression.

First, to test whether the two methods yield systematically different posture scores, we performed paired-samples comparisons on (i) the pooled session-level scores across all observations ($n = 40$) and (ii) participant-specific paired scores (each $n = 8$) to examine consistency of the between-method difference within individuals. For each paired comparison, we report the mean difference with a 95% confidence interval, the paired t statistic and two-tailed p-value, and the standardized effect size (Cohen's d based on the SD of paired differences). Table 3 summarizes the paired-test results.

Paired comparisons were conducted to examine between-method differences in posture assessment scores between the heuristic-based and video-based methods. Overall, across all 40 sessions, the video-based method yielded significantly higher scores than the heuristic-based method (heuristic: 89.38 ± 5.21; video: 97.38 ± 2.99), with a mean difference of 8.00 points (95% CI: 6.47–9.53; t (39) = 10.591, $p < 0.001$). The associated effect size indicated a large paired difference (Cohen's d = 1.675, 95% CI: 1.187–2.153).

Participant-level analyses showed the same trend. For P1–P4, the video-based scores were significantly higher than the heuristic-based scores, with mean differences of 8.75 points for P1 (95% CI: 5.79–11.71; t (7) = 7.000, $p < 0.001$; Cohen's d = 2.475), 10.63 points for P2 (95% CI: 7.95–13.30; t (7) = 9.379, $p < 0.001$; Cohen's d = 3.316), 8.13 points for P3 (95% CI: 4.30–11.95; t (7) = 5.017, $p = 0.002$; d = 1.774), and 8.75 points for P4 (95% CI: 3.88–13.62; t (7) = 4.249, $p = 0.004$; d = 1.502). For P5, the difference remained statistically significant but smaller in magnitude (mean difference: 3.75, 95% CI: 0.04–7.46; t (7) = 2.393, $p = 0.048$; d = 0.846). Collectively, these results indicate that the heuristic-based method systematically produced lower posture assessment scores than the video-based method across sessions and participants.

Table 3. Paired comparison of posture assessment scores between heuristic-based and video-based methods.

Analysis Unit	*n*	Heuristic S. *	Video S. *	Mean Difference (Video-Heuristic) [95% CI]	*t* (df)	*p* (Two-Tailed)	Cohen's d [95% CI]
Overall	40	89.38 ± 5.15	97.38 ± 2.99	8.00 [6.47–9.53]	10.591 (39)	<0.001	1.675 [1.187–2.153]
P1	8	85.00 ± 2.67	93.75 ± 2.31	8.75 [5.79–11.71]	7.000 (7)	<0.001	2.475 [1.014–3.902]
P2	8	88.13 ± 3.72	98.75 ± 2.31	10.63 [7.95–13.30]	9.379 (7)	<0.001	3.316 [1.473–5.136]
P3	8	90.00 ± 5.35	98.13 ± 2.59	8.13 [4.30–11.95]	5.017 (7)	0.002	1.774 [0.610–2.896]
P4	8	90.63 ± 6.23	99.38 ± 1.77	8.75 [3.88–13.62]	4.249 (7)	0.004	1.502 [0.444–2.516]
P5	8	93.13 ± 4.58	96.88 ± 2.59	3.75 [0.04–7.46]	2.393 (7)	0.048	0.846 [0.007–1.643]

* S. = Score.

Second, to test whether the two methods yield systematically different warning counts (discrete outcomes), we conducted paired comparisons using nonparametric tests to account for the paired structure of the observations. Specifically, we applied the Wilcoxon signed-rank test to (i) pooled session-level paired warning counts across all observations ($n = 40$) and (ii) participant-specific paired warning counts (each $n = 8$) to examine the consistency of between-method differences within individuals (P1–P5). For each comparison, we report the standardized test statistic (Z), the two-tailed *p*-value, and an effect size estimate computed as $r = |Z| / \sqrt{N}$. Table 4 summarizes the paired nonparametric test results.

Table 4. Paired nonparametric comparison of warning counts between heuristic-based and video-based methods.

Analysis Unit	*n*	Test	Z	*p* (Two-Tailed)	Effect Size *r* *
Overall	40	Wilcoxon signed-rank Sign test	−5.315 −5.918	<0.001 <0.001	0.841 -
P1	8	Wilcoxon signed-rank Sign test	−2.527 −2.475	0.012 0.008	0.893 -
P2	8	Wilcoxon signed-rank Sign test	−2.536 −2.475	0.011 0.008	0.897 -
P3	8	Wilcoxon signed-rank Sign test	−2.536 −2.475	0.011 0.008	0.897 -
P4	8	Wilcoxon signed-rank Sign test	−2.533 −2.475	0.011 0.008	0.895 -
P5	8	Wilcoxon signed-rank Sign test	−2.032 −1.789	0.042 0.063	0.718 -

* $|Z| / \sqrt{N}$.

Paired nonparametric analyses showed significant between-method differences in warning counts. At the pooled session level ($n = 40$), both the Wilcoxon signed-rank test ($Z = -5.315$, $p < 0.001$; $r = 0.841$) and the sign test ($Z = -5.918$, $p < 0.001$) indicated a systematic difference between the two methods. Participant-specific analyses (each $n = 8$) showed consistent results for P1–P4 (Wilcoxon $p = 0.011$–0.012; $r = 0.893$–0.897; sign test

$p = 0.008$). For P5, the Wilcoxon test remained significant ($Z = -2.032$, $p = 0.042$; $r = 0.718$), whereas the sign test was not significant ($p = 0.063$), suggesting that the direction of paired differences was less uniform for this participant despite an overall paired shift.

Finally, agreement between the video-based and heuristic-based posture scores was examined using Bland–Altman analysis. For each paired observation, we calculated the mean score of the two methods and the score difference between the two methods (video-heuristic). We then estimated the mean bias and the 95% limits of agreement (LoA; mean bias ± 1.96 SD of the differences) and visualized them on a Bland–Altman plot. To assess proportional bias, we performed a linear regression of the paired differences on the paired means; a significant slope was interpreted as evidence that the between-method difference varies with the magnitude of the measurement. Figure 8 shows the Bland–Altman plot, and Table 5 summarizes the proportional-bias regression results.

Table 5. Proportional-bias regression for Bland–Altman differences (video-heuristic) as a function of paired means.

Outcome	Predictor	B (SE) ***	95% CI for B	*t*	*p*	R^2 (Adj R^2)
V-H D *	V-H M **	−0.737 (0.185)	(−1.112)–(−0.362)	−3.977	<0.001	0.294 (0.275)

* Video-Heuristic score Difference, ** Video-Heuristic score Mean; *** B, regression coefficient; SE, standard error.

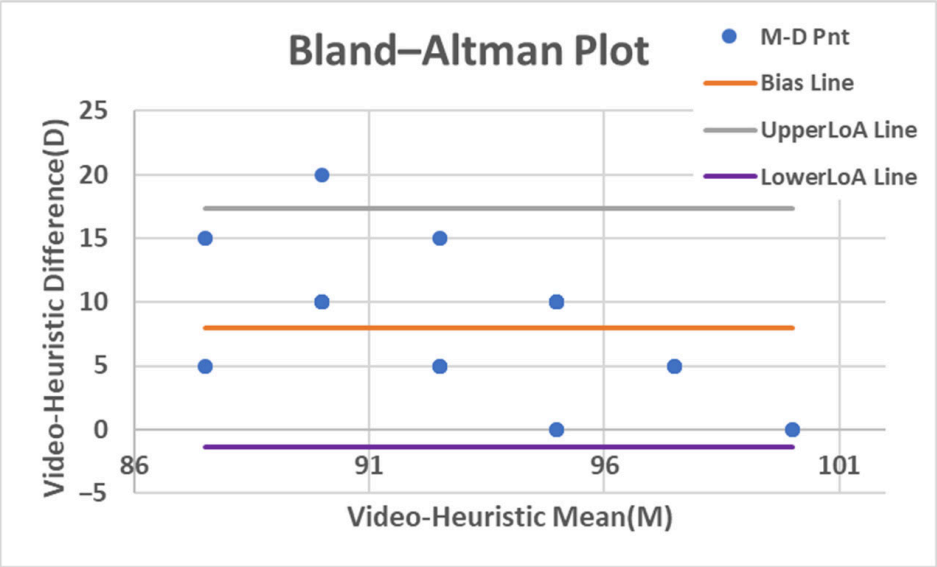


Figure 8. Bland–Altman plot of agreement between heuristic-based and video-based posture scores; Some points overlap because posture scores were recorded in discrete 5-point increments, resulting in identical mean–difference coordinates for multiple paired observations.

Agreement between the video-based and heuristic-based posture scores was evaluated using Bland–Altman analysis. For each paired observation, we computed the mean score of the two methods (M) and the paired difference (D = video-heuristic). As shown in Figure 8, the mean bias was 8.00 points, with 95% limits of agreement ranging from −1.36 to 17.36 points (bias ± 1.96 SD of the differences). To examine proportional bias, we regressed the paired differences on the paired means. The regression indicated a significant negative association ($B = -0.737$, $SE = 0.185$, 95% CI [−1.112, −0.362], $t = -3.977$, $p < 0.001$; $R^2 = 0.294$, adjusted $R^2 = 0.275$), suggesting that the between-method difference decreases as the overall score magnitude increases.

4. Discussion

We compared a video-based assessment with a motion-tracking-based heuristic assessment across eight clinical-task states (four procedure types combined with maxillary/mandibular conditions) using five dental residents. At the pooled session level, the heuristic consistently produced lower posture scores and higher warning counts than the video-based assessment, and this direction was largely consistent within individuals. This discrepancy should be interpreted as a consequence of the fundamentally different observation and decision structures of the two approaches, rather than as direct evidence of poorer operator performance. Bland–Altman analysis further showed a positive mean bias (video-heuristic) of 8.00 points with relatively wide 95% limits of agreement (−1.36 to 17.36 points), indicating that the two methods are not interchangeable at the individual-session level.

Rater-based judgments have structural limitations in quantitatively capturing continuous temporal changes. Decision sensitivity may vary for the same posture depending on observation conditions and rater interpretation of the criteria [72]. In addition, video-based assessment must infer 3D posture from 2D recordings and is therefore susceptible to camera viewpoint, limited field of view, and occlusion [73]. Recent studies have reported computer vision- and AI-assisted ergonomic risk assessment tools in real-world environments, including manufacturing and dentistry, highlighting both feasibility and context-dependent limitations [74,75]. Moreover, raters may under-reflect minor deviations that do not meaningfully affect task performance. Accordingly, future studies should explicitly ensure and report rater reliability by using standardized scoring rubrics and rater training, performing independent ratings by multiple specialists, and quantifying inter- and intra-rater agreement (e.g., ICC and/or weighted kappa) before consensus scoring is used as a reference.

By contrast, the heuristic assessment continuously computes distance, angle, and orientation from 3D joint coordinates captured by the motion-tracking device together with reference points from the virtual patient model. Adequacy is determined by comparing these measurements with predefined threshold intervals. Under this structure, brief threshold violations can accumulate as warnings and score reductions. In the present study, transition segments were not explicitly detected, labeled, or excluded; thus, short-lived posture readjustments and task transitions may have been counted as violations, inflating warning counts and contributing to the observed between-method discrepancy. Sensor-based coordinates may also include jitter, drift, and coordinate-frame alignment errors, which can increase sensitivity near threshold boundaries and further increase warnings [76,77]. These factors provide a plausible explanation for the higher warnings and lower scores produced by the heuristic method. Consistently, proportional-bias regression showed a significant negative slope, suggesting that disagreement decreases as overall score magnitude increases and may be larger in lower-scoring sessions.

Several session- and participant-level patterns are also noteworthy. The heuristic method yielded a broader distribution of warning and penalty counts, whereas video-based ratings were concentrated at relatively low penalty counts. This suggests that the heuristic framework is more sensitive to brief or repeated threshold violations across sessions, rather than differing only by a mean offset. Although the between-method direction was consistent overall, the magnitude of disagreement varied across participants and sessions. This supports the interpretation that disagreement is condition-dependent and may be influenced by transition frequency, task-specific movement characteristics, and threshold sensitivity near decision boundaries.

A key advantage of the heuristic framework is that it operationalizes posture criteria into measurable variables (e.g., distance, angle, orientation, and clock position) and automatically generates item-level decision logs. This enables large-scale collection of

time-series posture features and provides structured outputs for data-driven model development. The heuristic can therefore serve as an interpretable baseline and a practical mechanism for generating initial labels for model training, while expert annotations can be incorporated selectively to improve validity [78,79]. In addition, continuous quantitative signals may support calibration strategies (e.g., threshold tuning or bias correction) to improve agreement with practical reference assessments.

Although the present implementation targets dental posture training in a VR operatory, the pipeline is not inherently dentistry-specific. Its central contribution is a general method to operationalize domain guidelines into measurable kinematic variables and convert continuous time-series measurements into interpretable criterion-level decision logs and scores. By substituting the criterion library and reference geometry for a given workspace, the same sensing–feature–decision framework can be adapted to other procedural training contexts that require sustained ergonomics and constrained working postures. Future studies should evaluate cross-domain transferability by reparameterizing thresholds and reference points and verifying agreement against domain-specific reference assessments.

Importantly, these findings indicate comparative feasibility for objective posture quantification and assessment support, not improved clinical performance, reduced ergonomic risk, or long-term posture adoption. The system assessed adherence to predefined posture criteria and compared outputs with expert video-based ratings. However, procedural accuracy, treatment quality, efficiency, patient outcomes, and sustained ergonomic behavior in real clinical practice were not evaluated. Therefore, the educational value of the system should be framed as a supportive tool for training and assessment rather than a validated solution for long-term behavioral change.

Future work will develop AI-based posture assessment models using larger samples across broader proficiency levels [72]. We will export stabilized time-series signals (raw kinematic measurements after the current filtering pipeline) and construct windowed datasets aligned to clinical-task states, using expert-based assessments as reference labels. We will benchmark simple machine-learning baselines to identify informative feature sets and prediction targets (e.g., criterion-level violations and/or session-level posture scores), and then evaluate advanced temporal models that capture motion dynamics (e.g., CNN + LSTM) to improve agreement and generalization. To better handle transition-heavy segments, we will annotate or detect transition periods and evaluate models with and without these segments, together with temporal aggregation rules and temporal constraints (e.g., dwell-time rules), to reduce false positives from brief posture readjustments. We will also implement a real-time corrective feedback module to translate posture scores and violations into actionable guidance. In parallel, usability-oriented studies with dental students will evaluate learnability, perceived workload, and educational acceptability of the VR posture training and feedback workflow.

From a technical perspective, this study is limited by the small sample size and by the possibility that the video-based reference assessment was influenced by observation conditions and rater judgment. The small cohort limits statistical power and generalizability and constrains subgroup analyses across procedures and proficiency levels. Future work will recruit larger and more diverse cohorts to validate agreement patterns and reduce uncertainty in the estimated limits of agreement. The observed limits of agreement and proportional bias further suggest that agreement varies across sessions; thus, future studies should verify these patterns using more diverse tasks, participants, and reference standards.

From an educational perspective, the realism of the VR training environment and the transferability of posture behaviors acquired in VR to real clinical settings were not directly validated. Real clinical environments involve additional constraints (e.g., patient variability, assistant interaction, instrument-handling complexity, and workspace interruptions) that

may affect posture and assessment outcomes. Longitudinal posture improvement across repeated training sessions was also not assessed. Future studies should therefore evaluate skill transfer and retention using larger cohorts in real clinical settings or validation environments that incorporate key real-clinical elements.

5. Conclusions

This study compared video-based and motion-tracking-based heuristic posture assessments across eight clinical-task states performed by five dental residents. The heuristic approach tended to yield lower mean scores and more penalty occurrences than the video-based assessment; however, this discrepancy is more plausibly explained by differences in observation and decision structures (continuous threshold-based evaluation, transition segments, and sensor variability) rather than a deterioration in actual performance. Despite these differences, the proposed heuristic framework provides a practical means to quantitatively record posture and deliver an assessment during dental clinical practice. Moreover, by translating posture criteria into measurable variables and generating item-level decision logs, it offers a foundation for developing and validating AI-based posture assessment models. Limitations include the small sample size and potential bias in video-based ratings; future work will extend validation with larger and more diverse cohorts and investigate strategies (e.g., dwell-time rules) to reduce misclassification during transitional movements.

Author Contributions: Conceptualization, S.-Y.M. and J.-S.K.; methodology, S.-Y.M. and J.-S.K.; software, K.-W.K. and J.-S.K.; validation, J.-S.K. and H.-J.K.; formal analysis, J.-S.K.; investigation, J.-S.K. and H.-J.K.; resources, H.-J.K.; data curation, S.-Y.M. and H.-J.K.; writing—original draft preparation, J.-S.K.; writing—review and editing, S.-Y.M. and J.-S.K.; visualization, J.-S.K. and K.-W.K.; supervision, S.-Y.M.; project administration, J.-S.K.; funding acquisition, J.-S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government [Ministry of Science and Information and Communication Technology (MIST)] under Grant RS-2023-00214613.

Institutional Review Board Statement: Ethical review and approval were waived for this study in accordance with Article 15(2) of the Bioethics and Safety Act of Korea and Article 13(1)(3) of the Enforcement Rule of the Bioethics and Safety Act, because this study used existing data or documents from research participants and did not collect or record personally identifiable information.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Prior to participation, the study purpose and procedure were explained, and participants were informed that participation was voluntary and that no personally identifiable information would be collected (data were analyzed in anonymized form).

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kuper, H. Industry 4.0: Changes in work organization and qualification requirements—Challenges for academic and vocational education. *Entrep. Educ.* **2020**, *3*, 119–131. [CrossRef]
2. Rikala, P.; Braun, G.; Järvinen, M.; Stahre, J.; Hämäläinen, R. Understanding and measuring skill gaps in Industry 4.0—A review. *Technol. Forecast. Soc. Change* **2024**, *201*, 123206. [CrossRef]
3. World Economic Forum. The Future of Jobs Report 2025. 2025. Available online: https://www.weforum.org/publications/the-future-of-jobs-report-2025/in-full/ (accessed on 27 January 2026).

4. Strojny, P.; Dużmańska-Misiarczyk, N. Measuring the effectiveness of virtual training: A systematic review. *Comput. Educ. X Real.* **2023**, *2*, 100006. [CrossRef]
5. Scorgie, D.; Feng, Z.; Paes, D.; Parisi, F.; Yiu, T.W.; Lovreglio, R. Virtual reality for safety training: A systematic literature review and meta-analysis. *Saf. Sci.* **2024**, *171*, 106372. [CrossRef]
6. Angel-Urdinola, D.F.; Castillo-Castro, C.; Hoyos, A. *Meta-Analysis Assessing the Effects of Virtual Reality Training on Student Learning and Skills Development*; World Bank: Washington, DC, USA, 2021.
7. Howard, M.C.; Gutworth, M.B.; Jacobs, R.R. A meta-analysis of virtual reality training programs. *Comput. Hum. Behav.* **2021**, *121*, 106808. [CrossRef]
8. Bödding, R.; Schriek, S.A.; Maier, G.W. A systematic review and meta-analysis of mixed reality in vocational education and training: Examining behavioral, cognitive, and affective training outcomes and possible moderators. *Virtual Real.* **2025**, *29*, 44. [CrossRef]
9. Thomann, H.; Zimmermann, J.; Deutscher, V. How effective is immersive VR for vocational education? Analyzing knowledge gains and motivational effects. *Comput. Educ.* **2024**, *220*, 105127. [CrossRef]
10. Kablitz, D. Bridging Theory and Practice with Immersive Virtual Reality: A Study on Transfer Facilitation in VET. *Educ. Sci.* **2025**, *15*, 959. [CrossRef]
11. Shin, K.S.; Kim, J.C.; Cho, K.W.; Cho, W.I. Improving VR Welding Simulator Tracking Accuracy Through IMU-SLAM Fusion. *Electronics* **2025**, *14*, 4693. [CrossRef]
12. Long, Y.; Zhang, X.; Zeng, X. Application and effect analysis of virtual reality technology in vocational education practical training. *Educ. Inf. Technol.* **2025**, *30*, 9755–9786. [CrossRef]
13. Koçak, N.F.; Saygın, A.; Türk, F. Utilizing VR Technology in Foundational Welding Skill Development. *Appl. Sci.* **2025**, *15*, 12331. [CrossRef]
14. Michalak, D.; Rozmus, M.; Tokarczyk, J.; Szewerda, K. Portable VR Welding Simulator. *Appl. Sci.* **2024**, *14*, 7687. [CrossRef]
15. Checa, D.; Saucedo-Dorantes, J.J.; Osornio-Rios, R.A.; Antonino-Daviu, J.A.; Bustillo, A. Virtual reality training application for the condition-based maintenance of induction motors. *Appl. Sci.* **2022**, *12*, 414. [CrossRef]
16. Gómez-Cambronero, Á.; Miralles, I.; Tonda, A.; Remolar, I. Immersive virtual-reality system for aircraft maintenance education: A case study. *Appl. Sci.* **2023**, *13*, 5043. [CrossRef]
17. Li, S. Design and development of aviation aircraft maintenance training platform based on VR technology. *Procedia Comput. Sci.* **2023**, *228*, 898–906. [CrossRef]
18. Sun, Y.; Yang, Q.; Chen, J.; Li, H.; Wang, X.; Xu, D. Towards a multi-scenario immersive virtual reality system for maintenance inspection training of hot strip rolling mills. *Virtual Real.* **2025**, *29*, 86. [CrossRef]
19. Zhong, M.; Zhou, Y. Virtual-reality system for elevator maintenance education: Design, implementation and evaluation. *Eng. Rep.* **2024**, *6*, e12873. [CrossRef]
20. Wolfartsberger, J.; Zimmermann, R.; Obermeier, G.; Niedermayr, D. Analyzing the potential of virtual reality-supported training for industrial assembly tasks. *Comput. Ind.* **2023**, *147*, 103838. [CrossRef]
21. Di Pasquale, V.; Cutolo, P.; Esposito, C.; Franco, B.; Iannone, R.; Miranda, S. Virtual Reality for Training in Assembly and Disassembly Tasks: A Systematic Literature Review. *Machines* **2024**, *12*, 528. [CrossRef]
22. Maddipatla, Y.; Tian, S.; Liang, X.; Zheng, M.; Li, B. VR Co-Lab: A Virtual Reality Platform for Human–Robot Disassembly Training and Synthetic Data Generation. *Machines* **2025**, *13*, 239. [CrossRef]
23. Seo, S.; Park, H.; Koo, C. Impact of interactive learning elements on personal learning performance in immersive virtual reality for construction safety training. *Expert. Syst. Appl.* **2024**, *251*, 124099. [CrossRef]
24. Żuchowicz, P.; Lewczuk, K. Leveraging Immersive Technologies for Safety Evaluation in Forklift Operations. *Appl. Sci.* **2025**, *15*, 11048. [CrossRef]
25. Dianatfar, M.; Järvenpää, E.; Siltala, N.; Lanz, M. Template concept for VR environments: A case study in VR-based safety training for human–robot collaboration. *Robot. Comput.-Integr. Manuf.* **2025**, *94*, 102973. [CrossRef]
26. Pottle, J. Virtual reality and the transformation of medical education. *Future Healthc. J.* **2019**, *6*, 181–185. [CrossRef] [PubMed]
27. Kennedy, G.A.; Pedram, S.; Sanzone, S. Improving safety outcomes through medical error reduction via virtual reality-based clinical skills training. *Saf. Sci.* **2023**, *165*, 106200. [CrossRef]
28. Oxford Medical Simulation. Available online: https://oxfordmedicalsimulation.com (accessed on 1 November 2024).
29. Falah, J.; Khan, S.; Alfalah, T.; Alfalah, S.F.M.; Chan, W.; Harrison, D.K.; Charissis, V. Virtual Reality medical training system for anatomy education. In Proceedings of the 2014 Science and Information Conference, London, UK, 27–29 August 2014; pp. 752–758. [CrossRef]
30. Schild, J.; Misztal, S.; Roth, B.; Flock, L.; Luiz, T.; Lerner, D.; Herkersdorf, M.; Weaner, K.; Neuberaer, M.; Franke, A.; et al. Ap-plying multi-user virtual reality to collaborative medical training. In Proceedings of the 2018 IEEE Conference on Virtual Real-ity and 3D User Interfaces (VR), Tuebingen/Reutlingen, Germany, 18–22 March 2018; pp. 775–776. [CrossRef]

31. Ayoub, A.; Pulijala, Y. The application of virtual reality and augmented reality in Oral & Maxillofacial Surgery. *BMC Oral Health* **2019**, *19*, 238. [CrossRef] [PubMed]
32. Oyekunle, D.; Matthew, U.O.; Waliu, A.O.; Fatai, L.O. Healthcare applications of Augmented Reality (AR) and Virtual Reality (VR) simulation in clinical ceducation. *J. Clin. Images Med. Case Rep.* **2024**, *5*, 3141. [CrossRef]
33. Wang, N.; Yang, S.; Gao, Q.; Jin, X. Immersive Teaching Using Virtual Reality Technology to Improve Ophthalmic Surgical Skills for Medical Postgraduate Students. *Postgrad. Med.* **2024**, *136*, 487–495. [CrossRef] [PubMed]
34. Antaki, F.; Doucet, C.; Milad, D.; Giguère, C.-É.; Ozell, B.; Hammamji, K. Democratizing Vitreoretinal Surgery Training with a Portable and Affordable Virtual Reality Simulator in the Metaverse. *Transl. Vis. Sci. Technol.* **2024**, *13*, 5. [CrossRef] [PubMed]
35. Konovalov, A.; Okishev, D.; Pilipenko, Y.; Grebenev, F.; Artemyev, A.; Mamedbekova, G.; Eliava, S. Application of Virtual Reality for Developing a Ventriculostomy Simulator: Concept and Initial Results. *Surg. Neurol. Int.* **2025**, *16*, 525. [CrossRef] [PubMed]
36. Stevanie, C.; Ariestiana, Y.Y.; Anshar, M.; Sukotjo, C.; Boffano, P.; Forouzanfar, T.; Kurniawan, S.H.; Ruslin, M. Immersive Reality Surgical Training for Le Fort I Orthognathic Surgery: Initial Results of a Randomized Feasibility Study. *J. Cranio-Maxillofac. Surg.* **2025**, *53*, 1009–1017. [CrossRef] [PubMed]
37. McKinney, B.; Dbeis, A.; Lamb, A.; Frousiakis, P.; Sweet, S. Virtual Reality Training in Unicompartmental Knee Arthroplasty: A Randomized, Blinded Trial. *J. Surg. Educ.* **2022**, *79*, 1526–1535. [CrossRef] [PubMed]
38. Andersen, N.L.; Jensen, R.O.; Konge, L.; Laursen, C.B.; Falster, C.; Jacobsen, N.; Elhakim, M.T.; Bojsen, J.A.; Riishede, M.; Franse, M.L.; et al. Immersive Virtual Reality in Basic Point-of-Care Ultrasound Training: A Randomized Controlled Trial. *Ultrasound Med. Biol.* **2023**, *49*, 178–185. [CrossRef] [PubMed]
39. He, Q.; Li, J.; Xiao, Y.; Tian, J.; Mao, N.; Lu, S.; Zhao, Y. An Immersive Mixed Reality Surgical Self-Training System for Precise Pedicle Screw Placement: A Randomized Controlled Trial. *BMC Med. Educ.* **2025**, *25*, 1360. [CrossRef] [PubMed]
40. Jain, S.; Barua Chowdhury, B.D.; Mosier, J.M.; Subbian, V.; Hughes, K.; Son, Y.-J. Design and Development of an Integrated Virtual Reality (VR)-Based Training System for Difficult Airway Management. *IEEE J. Transl. Eng. Health Med.* **2025**, *13*, 49. [CrossRef] [PubMed]
41. Chuan, A.; Bogdanovych, A.; Moran, B.; Chowdhury, S.; Lim, Y.C.; Tran, M.T.; McLeod, G.A. Using Virtual Reality to Teach Ultrasound-Guided Needling Skills for Regional Anaesthesia: A Randomized Controlled Trial. *J. Clin. Anesth.* **2024**, *97*, 111535. [CrossRef] [PubMed]
42. Li, P.; Xu, B.; Zhang, X.; Fang, D.; Zhang, J. Design and development of a personalized virtual reality-based training system for vascular intervention surgery. *Comput. Methods Programs Biomed.* **2024**, *249*, 108142. [CrossRef] [PubMed]
43. Knobovitch, R.M.; Tokuno, J.; Botelho, F.; Fried, H.B.; Carver, T.E.; Fried, G.M. Virtual Reality Training Improves Procedural Skills in Mannequin-Based Simulation in Medical Students: A Pilot Randomized Controlled Trial. *Surg. Innov.* **2025**, *32*, 364–373. [CrossRef] [PubMed]
44. Liang, Y.; Huang, H.; Tan, Y.-B.; Li, T.; Huang, W.; Zhang, Q.-L.; Liu, Z.-W.; Kuang, M. Construction and Implementation of a Laparoscopic Skill Training Course Based on a Smartphone Application and Virtual Reality. *BMC Med. Educ.* **2024**, *24*, 1111. [CrossRef] [PubMed]
45. Chen, X.; Liao, P.; Liu, S.; Wang, L.; Zhao, Y.; Hu, J.; Zhang, Z. Effect of Virtual Reality Training to Enhance Laparoscopic Assistance Skills. *BMC Med. Educ.* **2024**, *24*, 29. [CrossRef] [PubMed]
46. Birrenbach, T.; Stuber, R.; Müller, C.E.; Sutter, P.M.; Hautz, W.E.; Exadaktylos, A.K.; Sauter, T.C. Virtual Reality Simulation to Enhance Advanced Trauma Life Support Trainings—A Randomized Controlled Trial. *BMC Med. Educ.* **2024**, *24*, 666. [CrossRef] [PubMed]
47. Zackoff, M.W.; Cruse, B.; Sahay, R.D.; Zhang, B.; Sosa, T.; Schwartz, J.; Depinet, H.; Geis, G.L. Multiuser Immersive Virtual Reality Simulation for Interprofessional Sepsis Recognition and Management. *J. Hosp. Med.* **2024**, *19*, 185–192. [CrossRef] [PubMed]
48. Zhang, N.; Ye, G.; Yang, C.; Zeng, P.; Gong, T.; Tao, L.; Zheng, Y.; Liu, Y. Benefits of Virtual Reality Training for Cardiopulmonary Resuscitation Skill Acquisition and Maintenance. *Prehosp. Emerg. Care* **2025**, *29*, 843–849. [CrossRef] [PubMed]
49. Siddique, H.; Abbas, A.; Abdul-Hamid, A. Effectiveness of Virtual Reality Simulation for Tracheostomy Education Among Healthcare Professionals. *Cureus* **2025**, *17*, e97083. [CrossRef] [PubMed]
50. Rahman, O.F.; Kunze, K.N.; Yao, K.; Kwiecien, S.Y.; Ranawat, A.S.; Banffy, M.B.; Kelly, B.T.; Galano, G.J. Hip Arthroscopy Simulator Training with Immersive Virtual Reality Has Similar Effectiveness to Nonimmersive Virtual Reality. *Arthroscopy* **2024**, *40*, 2840–2849.e3. [CrossRef] [PubMed]
51. Tashiro, Y.; Miyafuji, S.; Kojima, Y.; Kiyofuji, S.; Kin, T.; Igarashi, T.; Koike, H. MR Microsurgical Suture Training System with Level-Appropriate Support. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24), Honolulu, HI, USA, 11–16 May 2024*; Association for Computing Machinery: New York, NY, USA, 2024; pp. 1–19. [CrossRef]
52. McCloskey, K.; Turlip, R.; Ahmad, H.S.; Ghenbot, Y.G.; Chauhan, D.; Yoon, J.W. Virtual and augmented reality in spine surgery: A systematic review. *World Neurosurg.* **2023**, *173*, 96–107. [CrossRef] [PubMed]

53. Pagani, W.; Buysse, T.; Dua, K.S. Digital Platforms, Virtual Reality, and Augmented Reality in Gastrointestinal Endoscopy Training. *Clin. Endosc.* **2025**, *58*, 653–661. [CrossRef] [PubMed]
54. Henniger, D.; Engelke, M.; Kreiser, J.; Riemer, V.; Wierzba, E.; Dimitriadis, S.; Hann, A. Validation of the ViGaTu Immersive Virtual Reality Endoscopy Training System for Physicians and Nurses. *J. Gastrointestin. Liver Dis.* **2024**, *33*, 226–233. [CrossRef] [PubMed]
55. Yoon, H.; Lee, E.; Kim, C.J.; Shin, Y. Virtual Reality Simulation-Based Clinical Procedure Skills Training for Nursing College Students: A Quasi-Experimental Study. *Healthcare* **2024**, *12*, 1109. [CrossRef] [PubMed]
56. Jallad, S.T.; Işık, B. The Effectiveness of Immersive Virtual Reality Simulation as an Innovative Learning Strategy for Acquisition of Clinical Skills in Nursing Education: Experimental Design. *Games Health J.* **2025**, *14*, 110–118. [CrossRef] [PubMed]
57. Means, K.; Kleiman, K.; Ogdon, D.; Woodard, S. A review of virtual reality in radiology. *Curr. Probl. Diagn. Radiol.* **2024**, *53*, 17–21. [CrossRef] [PubMed]
58. O'Connor, M.; Rainford, L. The impact of 3D virtual reality radiography practice on student performance in clinical practice. *Radiography* **2023**, *29*, 159–164. [CrossRef] [PubMed]
59. Gohman, T.; Nisar, H.; Gupta, A.; Javed, M.J.; Rau, N. Development and Usability of a Virtual Reality Umbilical Venous Catheter Placement Simulator. *Int. J. Comput. Assist. Radiol. Surg.* **2024**, *19*, 881–889. [CrossRef] [PubMed]
60. Hsu, M.; Chang, Y.-C. Prediction of Dental Students' Pre-Clinical Crown Preparation Performances by a Virtual Reality Haptic-Based Dental Simulator. *J. Dent. Sci.* **2025**, *20*, 2301–2306. [CrossRef] [PubMed]
61. Rangarajan, K.; Davis, H.; Pucher, P.H. Systematic review of virtual haptics in surgical simulation: A valid educational tool? *J. Surg. Educ.* **2020**, *77*, 337–347. [CrossRef] [PubMed]
62. Arango, S.; Gorbaty, B.; Tomhave, N.; Shervheim, D.; Buyck, D.; Porter, S.T.; Laizzo, P.A.; Perry, T.E. A high-resolution virtual reality-based simulator to enhance perioperative echocardiography training. *J. Cardiothorac. Vasc. Anesth.* **2023**, *37*, 299–305. [CrossRef] [PubMed]
63. Kim, J.S.; Kim, K.W.; Kim, S.R.; Woo, T.G.; Chung, J.W.; Yang, S.W.; Moon, S.Y. An Immersive Virtual Reality Simulator for Echocardiography Examination. *Appl. Sci.* **2024**, *14*, 1272. [CrossRef]
64. Kim, K.W.; Kim, J.S.; Kim, H.J.; Moon, S.Y. Development and Evaluation of a Virtual Reality and Haptic-Integrated Clinical Simulation System for OSCE-Based Blood-Related Procedures. *IEEE Access* **2025**, *13*, 192947–192957. [CrossRef]
65. Ratushnyi, R.; Stakhanska, O. Ergonomic interventions and endodontic treatment outcomes: An analysis of dentist working posture and error rates. *Saudi Dent. J.* **2024**, *36*, 466–470. [CrossRef] [PubMed]
66. Gauthaman, J.; Ganesan, A. Ergonomic interventions in the dental setups and their role in preventing musculoskeletal disorders among dentists: A systematic review. *J. Lifestyle Med.* **2025**, *15*, 18. [CrossRef] [PubMed]
67. Jain, L.; Hegde, V.; Shetty, R.; Tandale, A.; Palekar, T.J.; Mehta, V.; Srindhi, S.R.; Negi, S. Prevalence of musculoskeletal disorders among dental professionals: An umbrella review. *Eur. J. Gen. Dent.* **2025**, *14*, 001–010. [CrossRef]
68. Kholinne, E.; Azalia, X.; Rahayu, E.P.; Anestessia, I.J.; Agil, N.; Muchtar. The prevalence and risk factors of musculoskeletal disorders among Indonesian dental professionals. *Front. Rehabil. Sci.* **2025**, *6*, 1513442. [CrossRef] [PubMed]
69. Gehrig, J.S.; Sroda, R.; Saccuzzo, D. *Fundamentals of Periodontal Instrumentation and Advanced Root Instrumentation*, 9th ed.; Jones & Bartlett Learning: Burlington, MA, USA, 2025; pp. 1–825.
70. Finkbeiner, B.L. *Four-Handed Dentistry: A Handbook of Clinical Application and Ergonomic Concepts*; Prentice Hall: Upper Saddle River, NJ, USA, 2001; pp. 1–64.
71. Korea Health Personnel Licensing Examination Institute. Available online: https://www.kuksiwon.or.kr/ (accessed on 3 January 2025).
72. Balogh, D.; Cui, X.; Mayer, M.; Koehncke, N.; Dueck, R.; Lang, A.E. Reliability and agreement during the Rapid Entire Body Assessment: Comparing rater expertise and artificial intelligence. *PLoS ONE* **2025**, *20*, e0323262. [CrossRef] [PubMed]
73. Song, B.; Wang, J.; Wang, X.; Zeng, T.; Li, D. Occlusion-aware and jitter-rejection 3D video real-time pose estimation for construction workers. *Autom. Constr.* **2025**, *171*, 106015. [CrossRef]
74. Agostinelli, T.; Generosi, A.; Ceccacci, S.; Mengoni, M. Validation of computer vision-based ergonomic risk assessment tools for real manufacturing environments. *Sci. Rep.* **2024**, *14*, 27785. [CrossRef] [PubMed]
75. Manohar, B.A.; Devaraj, J.; Maheswaran, C.; Pugalenthi, S. Performance Evaluation of Rapid Entire Body Assessment Using AI-Assisted Ergonomic Analysis in Dentistry. *Biomimetics* **2025**, *10*, 239. [CrossRef] [PubMed]
76. Yi, C.; Rho, S.; Wei, B.; Yang, C.; Ding, Z.; Chen, Z.; Jiang, F. Detecting and correcting IMU movements during joint angle estimation. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–14. [CrossRef]
77. Simon, S.; Meining, J.; Laurendi, L.; Berkefeld, T.; Dully, J.; Dindorf, C.; Fröhlich, M. 2D Pose Estimation vs. Inertial Measurement Unit-Based Motion Capture in Ergonomics: Assessing Postural Risk in Dental Assistants. *Bioengineering* **2025**, *12*, 403. [CrossRef] [PubMed]

78. Ratner, A.; Bach, S.H.; Ehrenberg, H.; Fries, J.; Wu, S.; Ré, C. Snorkel: Rapid training data creation with weak supervision. *VLDB J.* **2020**, *29*, 709–730. [CrossRef] [PubMed]
79. Gupta, P.; Nandakumar, S.; Gupta, M.; Panda, G. Data programming enabled weak supervised labeling for ECG time series. *Biomed. Signal Process. Control* **2024**, *87*, 105540. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.