Reflecting on Excellence: VR Simulation for Learning Indirect Vision in Complex Bi-Manual Tasks

Maximilian Kaluschke*
Computer Graphics & VR
University of Bremen

Rene Weller
Computer Graphics & VR
University of Bremen

Myat Su Yin
Faculty of ICT
Mahidol University

Benedikt W. Hosp
ZEISS Vision Science Lab
University of Tübingen

Farin Kulapichitr
Computer Graphics & VR
University of Bremen

Siriwan Suebnukarn
Faculty of Dentistry
Thammasat University

Peter Haddawy
Faculty of ICT
Mahidol University

Gabriel Zachmann
Computer Graphics & VR
University of Bremen

ABSTRACT

Indirect vision through a mirror, while bi-manually manipulating both the mirror and another tool is a relatively common way to perform operations in various types of surgery. However, learning such psychomotor skills requires extensive training; they are difficult to teach; and they can be quite costly, for instance, for dentistry schools. In order to study the effectiveness of VR simulators for learning these kinds of skills, we developed a simulator for training dental surgery procedures, which supports tracking of eye gaze and tool trajectories (mirror and drill), as well as automated outcome scoring. We carried out a pre-/post-test study in which 30 fifth-year dental students received six training sessions in the access opening stage of the root canal procedure using the simulator. In addition, six experts performed three trials using the simulator. The outcomes of drilling performed on realistic plastic teeth showed a significant learning effect due to the training sessions. Also, students with larger improvements in the simulator tended to improve more in the real-world tests. Analysis of the tracking data revealed novel relationships between several metrics w.r.t. eye gaze and mirror use, and performance and learning effectiveness: high rates of correct mirror placement during active drilling and high continuity of fixation on the tooth are associated with increased skills and increased learning effectiveness. Larger time allocation for tooth inspections using the mirror, i.e., indirect vision, and frequency of inspection are associated with increased learning effectiveness. Our findings suggest that eye tracking can provide valuable insights into student learning gains of bi-manual psychomotor skills, particularly in indirect vision environments.

Index Terms: K.3.1 [COMPUTERS AND EDUCATION]: Computer Uses in Education—Computer-assisted instruction (CAI); J.3 [LIFE AND MEDICAL SCIENCES]: Medical information systems; I.6.3 [SIMULATION AND MODELING]: Applications

1 INTRODUCTION

Dental surgery is a profession of precision, skill, and intricate coordination. Achieving proficiency in dentistry requires developing highly refined psychomotor skills for using a variety of dental instruments. Among the more difficult skills to develop is the proper use of indirect gaze through the dental mirror. The mirror is used to observe spaces during procedures that are not directly observable due to the angle or due to being occluded by parts of the patient's anatomy. For example, when working on upper molars, a dentist can normally not see the tooth directly due to the dentist's location relative to the patient, which is dictated by ergonomics. This is one of the most complicated psychomotor skills because it requires coordinated but only loosely coupled asymmetric bi-manual interaction with indirect vision while demanding high precision in a small space.

Figure 1: VR Simulator. A student is using our bi-manual VR dental simulator with two force-feedback devices.

Teaching dental skills is typically done through guided practice on mannequins fitted with realistic plastic teeth and jaws (phantom head) and later on real patients. Dental instructors can teach proper use of instruments such as the dental drill by observing the student's actions, providing verbal feedback and guidance, and demonstrating the correct use of the instrument. However, using the dental mirror does not lend itself well to such instruction since the observer (teacher or student) cannot directly observe what the person carrying out the procedure perceives through the mirror. This makes teaching proper mirror usage particularly challenging. Additionally, there is a high cost associated with expendable materials, such as plastic teeth and instruments. These expendable materials produce an estimated cost of around $32 000 per year for a medium-sized dental school of 500 students.

In recent years, technology has increasingly woven its way into dental training. Recent years have seen the proliferation of VR-based dental simulators due to enabling technological advancements, combined with concrete benefits of the approach [6, 20]. VR simulators offer high-fidelity simulations that can be configured so as to provide trainees with practice on various cases [1]. Also, their reusability could result in greatly lowered operational costs, which would lead to a quick amortization of the simulator cost. They also have the ability to record accurate data on individual performance, which provides the opportunity for trainees to practice independently and receive objective feedback [36]. A particularly important technological development is the integration of eye tracking into head-mounted displays, which permits objective analysis of user gaze. This pro-

---

[*] e-mail: mxkl@uni-bremen.de

vides the opportunity for the first time to gather objective data on mirror use during dental procedures.

This study uses eye tracking in an HMD-based dental VR simulator to provide a unique objective perspective on the use of indirect vision in learning complex asymmetric and loosely coupled bi-manual psychomotor skills. Specifically, we examine three questions:

1. What information on a dentist's skill level can be gained from mirror usage patterns?
2. How important is proper mirror usage for a successful learning trajectory of a dental student?
3. How does mirror usage in a VR simulator translate into performance on physical teeth?

Acquiring complex psychomotor skills as a key competence is not restricted to dentistry. For instance, laparoscopic surgery also requires highly precise and coordinated bi-manual handling of different instruments with indirect vision. We are confident that our findings have the potential to be a basis for reshaping future training methodologies in many domains.

2 RELATED WORK

In this section we present an overview of psychomotor training in dentistry, of VR dental simulators, and of how eye-tracking has been used in VR simulators to track user visual focus.

2.1 Psychomotor Training in Dentistry

Access cavity opening in endodontics is an important stage for instrument passage within root canals and to provide clear visualization, without which canals requiring treatment may be missed [4]. When dentists work within the narrow confines of root canals, the unaided human eye can only see a limited area up to the canal opening [27], thus dentists use dental mirrors, which require complex psychomotor skills [23, 5]. In general, indirect vision tasks are about 50% more challenging than direct vision tasks [23], yet dental schools often lack specific training [19], causing difficulties for students and dentists and, thus, they often decide to avoid using them or resort to using them by employing inadequate usage with improper posture in the process [21]. In addition, the necessity for clear vision in dentistry is not precisely defined, and there is no agreed-upon standard for the visual acuity required for an acceptable level of performance [25].

Recent approaches to addressing the need for training of mirror skills include the use of physical trainers and mirror skill training exercises in virtual simulation systems [24, 26]. Mirroprep is a physical indirect vision training device that consists of a square mirror and a replica of a dental drill with a pencil attached. Students practice by tracing figures on a piece of paper, while viewing through the mirror. A study by Rau et al. [28] recruited students at different stages of dental education and evaluated their performance in a drawing exercise with indirect vision using the Mirroprep. Their findings revealed that Mirroprep practice can enhance and develop mirror vision skills [28].

Perceptual learning and practice are needed to help dental students acquire muscle memory to develop mirror skills [23, 5]. To be effective, a practical training approach to mirror training should replicate the same orientation and visual reference points encountered during dental procedures. McClure and colleagues [24] introduced the Jumpstart Mirror Trainer which incorporates a rotatable jaw, handpiece-shaped pencil, and multiple interchangeable arches containing activities that replicate various dental procedures. The authors compared the effectiveness of the Jumpstart Mirror Trainer with the Mirroprep in teaching indirect motor skills. The results showed that activities with the Jumpstart Mirror Trainer led to significantly improved student scores in Class I cavity preparation on upper teeth compared to those using the Mirroprep and the control device.

2.2 VR Simulators

The Simodont virtual dental trainer (MOOG, NieuwVennep, The Netherlands) offers mirror exercises integrated with a functional dental mirror for indirect vision tasks. In a study by Chu et al. [3], seventy-two dental students were randomly split into control and the experimental groups. Both groups received training with the Simodont system and underwent the initial mirror operation examination. The experimental group practiced using Mirrosistant and the control group using traditional methods with dentognathic models, complete dentitions, dental mirror and manual instruction. Their findings showed that students preferred traditional indirect dental mirror training on plastic teeth with an actual handpiece to the use of the virtual simulation dental training system. The authors mentioned that the virtual simulation dental training system did not give a realistic sense of manipulation and did not mimic the narrow operating space of the mouth.

2.3 Eye Tracking in VR Simulators

Eye-tracking technology has revolutionized the way we understand human perception and cognitive processes. By capturing the subtle movements of the eyes, researchers and professionals can gain insights into attention, cognitive load, and decision-making processes. Eye tracking devices, such as the HTC Vive Pro Eye (VIVE) or XTAL from Vrgineers, have advanced significantly over the years. These systems offer precise gaze data, allowing for a detailed analysis of where an individual is looking at any given moment. The integration of eye tracking into VR simulators has opened up new avenues for research and training. When combined, VR and eye-tracking technologies can offer comprehensive insights into trainees' actions, intent, and focus [12]. For instance, VR has emerged as a pivotal tool in medical training, providing a risk-free environment for practice. The bridge to expertise in medical training through VR showcases the advancements in technology and its potential to shape future training methodologies. Eye tracking has been instrumental in distinguishing between experts and novices in various surgical domains [35, 34, 13]. There is also the idea to facilitate adaptive training by detecting a user's focus on specific areas during task execution [2]. In arthroscopic shoulder surgery, eye tracking metrics have been used to detect periods of confusion as students navigate the shoulder joint with the arthroscope [35]. The authors in [34] investigate how dental students' levels of expertise influence their clinical performance in terms of dwell time on each tooth location, total examination time, and perceived task load in the virtual dental lab. In their study, participants were divided into expertise groups and tasked with virtual simulations for dental caries detection and diagnosis. Their findings suggest that the level of expertise significantly affect the performance of dental examinations in all areas except the anterior maxillary teeth. Both total dwell time on the dental mirror and total examination time were considerably shorter for the high expertise group than for the medium and low expertise groups. Generally, as students gain experience with virtual reality simulations, the total dwell time needed for clinical examinations decreases [14], indicating the reduced time required for clinical performance. While eye tracking in VR offers numerous advantages, it also poses challenges [22]. The integration of both technologies requires careful calibration to ensure accurate data capture. Additionally, the immersive nature of VR can sometimes lead to discrepancies in gaze data. Understanding these limitations is crucial to harness the full potential of eye tracking in VR.

3 VIRTUAL SIMULATOR

We have developed a VR dental simulator with bi-manual haptic control, building upon the simulator presented in [17] (see Fig. 1), which is based on the Unreal Engine (UE) 4.27.2 (with the modifications from Sect. 3.3). It uses the HTC Vive Pro Eye HMD and two 3D Systems Touch™ haptic devices. At the core of the simulation

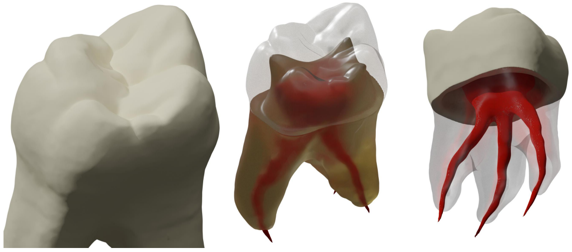

Figure 2: Anatomical tooth model. Our model of tooth #26, manually designed based on CT data. *Left:* The crown of the tooth, which consists of enamel. *Center:* The inner dentin layer. *Right:* The pulp chamber and roots.

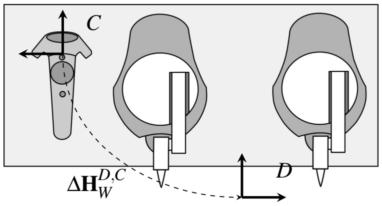

Figure 3: Haptic device setup. The registration of the haptic devices is simplified by mounting them, together with a VR controller, on predefined positions on a common surface.

are our algorithms for force feedback and material removal on the tooth, which run at 2000 Hz [15].

In this section, we give an overview of the simulator from a technical perspective, and provide details of the newly added features, as well as some of the technical challenges, including registration of VR space and haptic devices (Sect. 3.2), simulation of optical magnification (Sect. 3.3), and integration of eye tracking into UE (Sect. 3.4).

Other challenges of, arguably, greater import for the usability and effectiveness of the simulator, such as

- Convincing and stable force feedback,
- Fast and faithful material removal,
- Realistic and interactive material visualization,

were presented in an earlier publication [15].

3.1 Anatomical Tooth Model

We have manually designed a maxillary tooth model (see Fig. 2). We chose the maxillary tooth #26, as this would require the use of a dental mirror to have good viewing angles during tooth preparation. The tooth is divided into three layers: Enamel, dentin and pulp/root. For each layer, we constructed a closed surface mesh with Blender, while paying close attention to anatomical correctness, which was later also verified by expert dentists. The inner volume of each surface mesh is then approximated by a set of spheres, which are located on the inside of the surface mesh, similar to [15]. We tag each sphere based on the layer it belongs to. This tag is used during run-time to decide the material properties, such as visual color and roughness, and density. The average contact density is used to modulate the drilling speed, such that dentin is much easier to remove than enamel.

3.2 Haptic Device Registration

In order to maximize spatial presence and learning efficacy, there should be a perfect spatial match between the virtual tools and the real handles of the devices, i.e., perfect visuo-haptic synchronicity [7]. Without registration of the haptic devices into the virtual world, only the relative motion of the virtual tools would be correct, but their position in the virtual world would be more or less offset from where the user would expect them due to the user's kinaesthetic sense.

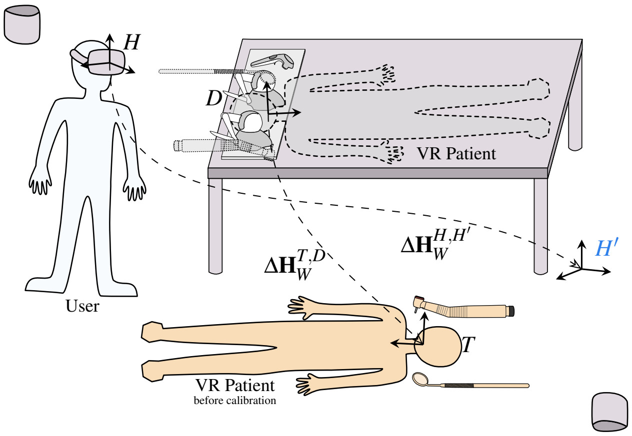

Figure 4: VR Registration. Transformations between coordinate systems are denoted in bold letters; they are explained in Section 3.2. The registration is done by moving the VR camera to H', such that the difference between physical HMD and physical device is the same as for virtual camera and virtual tool. The result is that the virtual tools and real haptic handles align in virtual and real space.

Our registration procedure aims to minimize the positional and rotational error between the real haptic device handles and the virtual haptic tools, which are "attached" to the haptic devices so as to follow their movements. To facilitate this, we have mounted the haptic devices on a rigid wooden board, together with a VR controller (see Fig. 3), thereby fixing the relative poses of both haptic devices and VR controller. Thus, the two haptic coordinate frames can be considered as one with the origin in the middle between them. We can move this origin with a fixed transformation into the virtual patient's mouth, such that the virtual tools are to the left and right of the virtual patient's mouth when the real haptic handles are held in a neutral position. The VR controller on the board serves as kind of an "extension" of the SteamVR tracking system to the haptic devices. What remains to be done is the transformation of the virtual camera such that offset between camera and virtual tools is exactly the same as the offset between the user's HMD and the haptic handles.

Let $\mathbf{H}_{W\leftarrow L}$ denote a transformation matrix that transforms points from a local space L to world space W (such a transform also gives rise to a pose). As explained previously, the goal of the registration is to adjust the camera pose $\mathbf{H}_{W\leftarrow H}$ to a new pose $\mathbf{H}_{W\leftarrow H'}$, such that the virtual tools occupy the same *visual* location as the haptic device handles' *physical* locations. By moving just the camera, we (i) avoid modifying the virtual scene, and (ii) we do not need to move any physical objects around. The delta transformation between poses H and H', given in world frame, is

$$\Delta\mathbf{H}_W^{H',H} = \mathbf{H}_{W\leftarrow H'}\,\mathbf{H}_{W\leftarrow H}^{-1} \qquad (1)$$

which can be computed and saved to disk once during registration, since it only needs to be updated when the setup changes.

In our simulator, the virtual tool pose, $\mathbf{H}_{W\leftarrow T}$, is set in the UE scene graph, which allows for convenient and intuitive adjustment of the device workspace center. The real haptic device pose, $\mathbf{H}_{W\leftarrow D}$, in UE world can be derived from the VR controller's pose, $\mathbf{H}_{W\leftarrow C}$, which is mounted next to the haptic device (see Fig. 3). Thereby, the offset $\Delta\mathbf{H}_W^{D,C}$ of the haptic device relative to the VR controller pose $\mathbf{H}_{W \leftarrow C}$ is now a constant offset, which needs to be registered manually, once. For our setup, we manually tuned this constant offset to be:

$$\Delta\mathbf{H}_W^{D,C} = \mathbf{T}(\,(37\,\mathrm{cm}, 26\,\mathrm{cm}, -7\,\mathrm{cm})^T\,)\,\mathbf{R}_z(-\frac{\pi}{2}) \tag{2}$$

The haptic device pose is then given by

$$\mathbf{H}_{W \leftarrow D} = \Delta\mathbf{H}_W^{D,C}\,\mathbf{H}_{W \leftarrow C} \tag{3}$$

Given $\mathbf{H}_{W \leftarrow T}$ and $\mathbf{H}_{W \leftarrow D}$, we can now calculate the adjusted camera pose $\mathbf{H}_{W \leftarrow H'}$ by offsetting $\mathbf{H}_{W \leftarrow H}$, using

$$\mathbf{H}_{W \leftarrow H'} = \Delta\mathbf{H}_W^{H',H}\,\mathbf{H}_{W \leftarrow H} \tag{4}$$

with

$$\Delta\mathbf{H}_W^{H',H} = \Delta\mathbf{H}_W^{T,D} = \mathbf{H}_{W \leftarrow T}\,\mathbf{H}_{W \leftarrow D}^{-1} \tag{5}$$

In UE this can be accomplished by giving all VR objects (HMD camera and controllers) a common parent node with a non-zero transformation $\Delta\mathbf{H}_W^{H',H}$. We can now show $\mathbf{H}_{H' \leftarrow T} = \mathbf{H}_{H \leftarrow D}$:

$$\begin{aligned}\mathbf{H}_{H' \leftarrow T} &= \mathbf{H}_{W \leftarrow H'}^{-1}\,\mathbf{H}_{W \leftarrow T} = (\Delta\mathbf{H}_W^{H',H}\,\mathbf{H}_{W \leftarrow H})^{-1}\,\mathbf{H}_{W \leftarrow T} \\ &= (\Delta\mathbf{H}_W^{T,D}\,\mathbf{H}_{W \leftarrow H})^{-1}\,\mathbf{H}_{W \leftarrow T} = \mathbf{H}_{W \leftarrow H}^{-1}\,(\Delta\mathbf{H}_W^{T,D})^{-1}\,\mathbf{H}_{W \leftarrow T} \\ &= \mathbf{H}_{W \leftarrow H}^{-1}\,\mathbf{H}_{W \leftarrow D} = \mathbf{H}_{H \leftarrow D}\end{aligned} \tag{6}$$

Therefore, the real device is (relative to the user), at the same location as the virtual tool is (relative to the VR camera). Thereby, they are physically located and visually rendered at the same location.

In practice, the setup procedure by SteamVR (or others) provide a world coordinate frame well aligned with the (real) floor; we also assume that the table top is parallel to the floor. So, we only need to consider translation, and rotation along z in $\Delta\mathbf{H}_W^{H',H}$, which simplifies the manual registration part.

3.3 Simulation of Optical Magnification in VR

In reality, dentists regularly make use of magnifying binoculars during their procedures. Therefore, we implemented a similar feature in our VR simulator that allows users to switch between different levels of optical magnification on the fly (1x, 2x, 4x, 8x, 16x). However, the UE does not allow developers to make any changes to the camera projection when the scene is rendered in an HMD. In most cases, this makes sense, in order to prevent motion sickness, e.g., by an incorrect field-of-view. However, we believe our case warrants an exception for optical magnification, for the following reasons

1. There is a clear physical meaning behind the magnification since dentists often use surgical binoculars.
2. The user is in control of the magnification. If they dislike it, or feel sick, they can disable it quickly.
3. The user is always seated, and during the procedure, only very small, and very controlled head movements are done.
4. The user can see the tooth and, in particular, the root canal more clearly, since most current HMD's do not offer enough resolution to render such details clearly.

We will see in the data later-on that most users make frequent use of the magnification.

To implement the optical magnification, we modified the source code of UE. The optical zoom of a factor m is implemented by manipulating UE's default stereo projection matrix $\mathbf{H}_{S \leftarrow W}$ (which transforms a 3D point in camera space to a 2D point in screen space) as follows

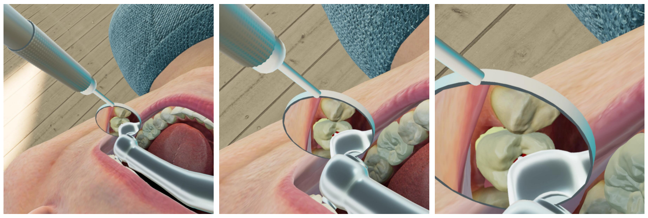

Figure 5: Optical magnification. Screen capture showing the same screen region at different magnification levels. When looking at the same screen position, the occupying object changes when the level of magnification changes. **Left:** 1x. **Center:** 2x. **Right:** 4x.

$$\mathbf{H}_{S \leftarrow W} \leftarrow \begin{bmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{H}_{S \leftarrow W} \tag{7}$$

We implemented this modification of UE in the function `GetStereoProjectionMatrix` of SteamVR and OpenXR. Thus, a point that would normally be projected onto $\mathbf{p}$ in screen space is instead projected onto $m\mathbf{p}$. This effectively zooms the screen image by m.

Additionally, we need to adjust the positions of both eyes, which are used in SteamVR and OpenXR's `GetRelativeEyePose`. In that function, the eye-to-head transformation is created from some translation $\vec{t}$ and a rotation. The distance of the eyes to the HMD origin needs to be scaled down by

$$\vec{t} \leftarrow \vec{t}\,\frac{1}{m} \tag{8}$$

where m depends on the magnification factor chosen by the user. Without this correction, our magnification would also incur an increased stereopsis, which needs to remain unchanged, in order to produce correct spatial impressions. The choice of m defaults to 1 and can be adjusted by the user by pressing the front or back button of the left 3D Systems Touch.

In our user study, novices spent on average 77 % and experts 60 % of their whole training time in the optically magnified mode. In fact, most participants (52 %) spent nearly all of their training time in zoomed mode (over 95 % of the time), with 2x being the most common setting. Two participants reported experiencing discomfort when using optical magnification.

3.4 Eye Tracking

The HTC VIVE Pro Eye has built-in sensors that track the user's eyes at a frequency of 120 Hz, with an accuracy of 0.5° to 1.1°, by hardware specifications. We use the SRanipal SDK to communicate with the eye sensors, which works very well. This SDK also provides a UE plugin; however, we have found two notable shortcomings of that plugin, which we explain in the following.

(i) Low Frequency of Sensor Readings The SRanipal UE plugin works synchronously with the rendering thread. Therefore, sensor updates are bound by the rendering performance of our application. In our case, we usually run between 60 Hz to 90 Hz, as we are rendering a demanding VR scene with geometry that is constantly updated at run-time. This means we could only record eye tracking data at around 75 Hz, as opposed to the advertised 120 Hz. Leube et al. [22] have shown that there is a significant decrease in saccade detection when going from 120 Hz to 60 Hz. In addition, the net frequencies will be even lower, as some frames need to be rejected because of incorrect sensor readings.

(ii) Gaze Origin Inaccuracy The SRanipal UE plugin provides a function to compute a raycast from the user's cyclops eye (the midpoint between the left and right eye) into the virtual scene, to determine which virtual 3D point the user is focusing on. However, upon inspecting the source code for this function, we found it incorrectly assumes the user's cyclops eye at $\vec{0}$ inside the camera's local frame during the raycast. Thus, the gaze origin is simply replaced by the camera position (the gaze direction is correctly transformed to the UE world). Obviously, a raycast from an incorrect origin will decrease the accuracy of the position the user fixates. This could be one of the reasons why several studies that examined the VIVE's eye-tracking accuracy found significant deviations from the advertised hardware accuracy: $1.71°$ by [29] and $4.16°$ by [31].

To alleviate both issues, we have implemented the sensor communication in a C++ library that we access from UE. Our C++ library is based on the SRanipal C++ SDK, but it allows for running the eye tracking in a separate thread that runs asynchronously to the game engine. Thus, we achieve the maximum tracking frequency of 120 Hz. Additionally, we corrected the incorrect gaze origin transformation from sensor space to UE world space. The gaze direction transformation needed to be adjusted to incorporate the changes to the stereo projection matrix described in Eq. 7 and Eq. 8.

Given a gaze origin $\mathbf{o}_S$ and direction $\vec{d}_S$ that are defined in the HMD's local sensor frame, S, we compute the gaze origin and direction $\mathbf{o}_W, \vec{d}_W$ in the world frame by first transforming them into camera space using

$$\mathbf{H}_{C\leftarrow S} = \mathbf{S}(l_W)\,\mathbf{S}(\tfrac{1}{m})\,\mathbf{S}(-1,1,1)\,\mathbf{R_z}(-\tfrac{\pi}{2})\,\mathbf{R_x}(\tfrac{\pi}{2})\,\mathbf{S}(0.001) \qquad (9)$$

where l_W is the UE world-to-meter property, which we set to 50, and $\mathbf{S}$ is a scaling transform. This transformation is constant since the sensor does not move relative to the camera.

In the case the user has switched on the magnification (i.e., dental loupe), the gaze direction $\vec{d}_C$ needs to be rotated to adjust for the different positions that objects assume on screen (see Fig. 5 as an example). We create a rotation $\Delta\mathbf{R}^{\vec{x},d}$ that rotates the gaze direction $\vec{d}_C$ towards $\vec{x} = (1,0,0)$ (in UE, this is the forward direction in camera space). We then compute the corrective rotation through spherical interpolation by $1-\frac{1}{m}$ of the rotation towards the 0-rotation

$$\Delta\mathbf{R}^{d',d} = \mathrm{slerp}(\Delta\mathbf{R}^{\vec{x},d}, \mathbf{0}, 1-\tfrac{1}{m}) \qquad (10)$$

where the direction is defined as: $\mathrm{slerp}(\mathbf{A},\mathbf{B},0) = \mathbf{A}$ and $\mathrm{slerp}(\mathbf{A},\mathbf{B},1) = \mathbf{B}$.

Finally we transform from camera space to world space by $\mathbf{H}_{W\leftarrow C}$ (this matrix is dynamic and can be retreived from the UE scene node). The total transformation then is

$$\mathbf{o}_W = \mathbf{H}_{W\leftarrow C}\,\mathbf{H}_{C\leftarrow S}\,\mathbf{o}_S \qquad (11)$$

$$\vec{d}_W = \mathbf{H}_{W\leftarrow C}\,\Delta\mathbf{R}^{d',d}\,\mathbf{H}_{C\leftarrow S}\,\vec{d}_S \qquad (12)$$

3.5 Measurement of Gaze Behavior

In order to analyze the gaze behavior of participants, we implemented a logging system that logs data of the user's gaze and the current simplified simulation state. The logs are then processed afterward to generate descriptive statistics on the behavior related to mirror placement, visual focus, etc.

We are, in particular, interested in two cases: (i) when the user is looking at the tooth or the bur, and (ii) when the user is inspecting the tooth. We will denote case (i) by the term *Vision*; we can detect this by casting the eye gaze ray into the scene and checking whether it hits the tooth or the bur geometry; in case it hits the mirror, we follow the reflected ray. This will also easily catch occlusions of the bur or the tooth by the handpiece, or incorrect placement of the mirror. Obviously, Vision is a desirable state for the full length of the procedure. Case (ii) is denoted by *Inspection* and can be detected by checking for a hit of the ray with the tooth while the bur is not removing material from the tooth.

We implemented the detection of the events Vision and Inspection, which are marked by breaks in temporal coherence, as follows:

- **Vision break:** if there was no continuous Vision (of either bur or tooth) during the previous second. Obviously, a state of Vision is beneficial, and no Vision with the bur removing material is to be minimized. Therefore, we track vision breaks overall, and vision breaks during drilling.

- **Inspection:** this event is recorded as soon as there was no material removal for three seconds and the user's gaze hits the tooth surface. Inspection time is ended once material removal started again or there is no Vision of the tooth for one second. Therefore, dentists are generally not inspecting, and make use of Inspection occasionally to inspect their progress more closely.

We keep track of count, frequency, and average length of Vision breaks and Inspections, to be used in the analyses later. Additionally, we will also refer to the overall ratio of unobstructed indirect vision during drilling in relation to the whole time drilling under the term "correct mirror pose in drilling". We have introduced these new metrics, as metrics presented in previous work are not suitable to this particular task. Previous metrics are either too low-level, due to being designed for machine learning [13, 35], or not suited to our use-case [34]. In our case, vision is always concentrated at the same focus points, just the correct indirection through the mirror and avoidance of visual obstruction between mirror and AOI influence the gaze.

Our eye gaze recognition was overall quite reliable. The eye sensors provided valid eyetracking data in 74.31 % of all measurements. However, a few participants had lower sensor validity, primarily due to participants not wearing the HMD properly. We decided to exclude those participants from the analyses in order to minimize noise in the eyetracking data.

In order to ensure the validity of the recordings of the eye tracking data, we conducted accuracy checks, at the beginning of each trial. This check consisted of presenting 3×3 red dots in order while instructing the participants to fixate them. These red dots were located on a plane located 30 cm in front of the camera. They deviated from the central viewing direction by $2.86°$ to $14.04°$, with $8.85°$ on average. The dot pattern was biased downwards by $5.71°$ to better sample the relevant viewing directions since dentists tend to look mostly downwards (relative to the central viewing direction) during the procedure. During the check, we recorded the median angle between the gaze ray and the ray towards the currently shown red dot. Each dot was shown for 3 s, of which we discarded the first second and the last 0.5 second to allow the participants enough time to change fixation targets. In our study, participants achieved an accuracy of $0.29°$ to $3.58°$, with on average $1.21°$. In order to further minimize noise, we excluded trials if their accuracy error exceeded $\varepsilon_a = 1.27°$. We derived this threshold, ε_a, by taking $1/3$ of the diameter of the top surface of the tooth at a distance of 25.5 cm. The rationale for this is that we want gaze rays near the middle of the tooth surface, perpendicular to it, to be correctly recognized by the eye tracking.

3.6 Scoring

We need to assess to skill level of the participants before and after the training in VR. Since we are interested in their skill level in real life, we do so by having participants perform the procedure on

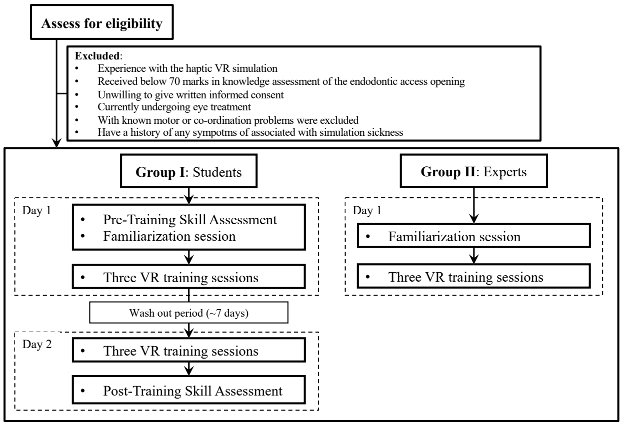

Figure 6: Study flowchart. Flowchart that shows the individual steps of the user study procedure.

plastic teeth just as they are used to in traditional training of dental students.

Both pre- and post-training plastic specimens were evaluated independently by two dental experts. The specimens were submitted to the evaluation experts anonymously, and they had no clue as to which specimens originated from before or after the training. The experts conducted the assessment using a standardized scoring system for root canal access opening, which has a scale from 0 to 15, with 0 indicating an optimal result devoid of errors. The cumulative score represents errors across five distinct tooth regions: the four walls of the opening and the pulp floor. Each region was scored as follows: 0 for the absence of errors, 1 for minor under-drilling, 2 for minor over-drilling, and 3 for significant over-drilling leading to structural instability or under-drilling rendering the orifice inaccessible. The experts' scoring results exhibited excellent reliability (intraclass correlation coefficient of 0.994, $\kappa = 0.817$). Owing to this strong agreement, we will utilize the mean of both ratings in the following analyses for interpretive clarity. To quantify the real learning gain, we consider the error difference Δe between pre- e_1 and post-training error e_2. Therefore a negative Δe represents a learning gain, whereas a positive Δe represents a performance decrease. Thereby, both outcomes, learning gain and error, consider lower numbers as the more desirable outcome.

To analyze participants' performance during training and the subsequent learning process, we collected their drilling outcomes in the form of a volumetric model. To judge the outcome error, we utilized a binary classification method, which we adapted from [18]. The general idea of this method is to classify the drilled and undrilled voxels against an ideal drilling outcome, leading to four classes (true positive, true negative, false positive, false negative), which can be used to score the outcome relative to the ideal. The different errors of over-drilling and under-drilling are given different weights and the error is scaled to roughly match the dental scoring system. We calculate the simulator learning gain analog to the real learning gain, with e_1 as the simulator score on trial #1 and e_2 as the simulator score on trial #6.

4 USER STUDY DESIGN

In this section, we present details on the design of the user study, including the demographics of participants and the individual steps during data collection. We have received ethical approval to conduct this user study from the Institutional Review Boards of Mahidol and Thammasat universities.

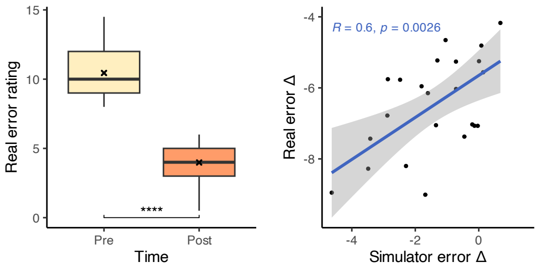

Figure 7: Learning gains. *Left*: The real-world performance of participating dental novices before and after training. ****: $p < .0001$. *Right*: Correlation of real and simulator learning gain (lower means higher error reduction).

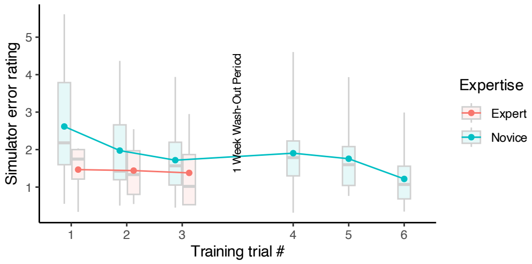

Figure 8: Simulator performance. The performance of novices and experts over the trial period. The error ratings are automatically generated.

4.1 Participants

We recruited 30 participants (7 male, 22 female, 1 undisclosed) and six experts (2 male, 3 female, 1 undisclosed) with different experience levels. All student participants were fifth-year dental students, with an average age of (26.83 ± 7.23) years. For the analysis, we excluded seven students and two experts due to problems during the training. The problems that warranted exclusion were either heat exhaustion from a malfunctioning AC, poor eye-tracking data validity, or sight problems from the HMD not fitting due to eyeglasses. The vast majority of participants usually wore glasses (73 %), though most could see well enough without them to proceed without or their glasses fitting inside the HMD.

4.2 Procedure

They were not admitted to the study if any of the following criteria were present: (i) had prior experience with the simulation, (ii) received below 70 % marks in knowledge assessment of endodontic cavity preparation, (iii) unwilling to give consent, (iv) have prior experience with VR systems. While the experts were asked to carry out three trials of the access opening procedure using the VR simulator, the task for the student participants was to perform access opening on the virtual tooth during the training session and on a plastic tooth (upper left molar; tooth number 26; http://www.nissin-dental.net/) in pre- and post-training assessment sessions. The plastic teeth resemble the feeling of drilling real teeth and are anatomically correct. A student's ability to perform the root canal access opening on such plastic teeth will predict with high reliability their ability to perform the task on real human teeth. Thus, using plastic teeth is the best option to assess real-world dental skills that is also ethically sound. Participants were briefly instructed on using the simulator, the experiment flow, and the requirements of the access opening.

As shown in the study flowchart (see Fig. 6), each participant's training took place on two separate days. The first day consisted of a briefing, pre-test, a familiarization session, and the first training session using the simulator, consisting of three trials. During the

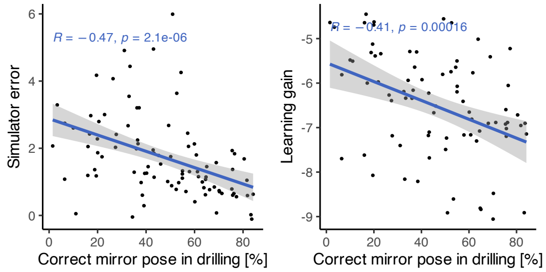

Figure 9: Correct mirror placement during drilling. *Left:* The relation between correct mirror placement (measured as a proportion of time drilling with correct vision relative to the time spent drilling) on simulator performance. *Right:* The relation between correct mirror placement and real-world learning gain.

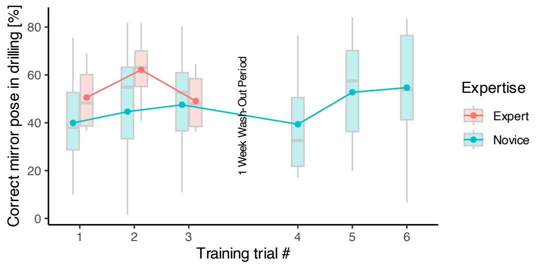

Figure 10: Mirror placement over time. The participants' correct mirror placement (measured as a proportion of time drilling with correct vision relative to the time spent drilling) over time, grouped by expertise.

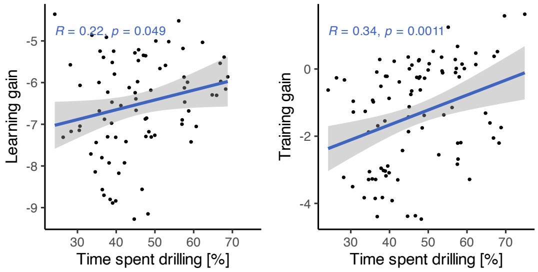

Figure 11: Time spent drilling & learning. *Left:* The relation between time spent drilling and real-world learning gain. *Right:* The relation between time spent drilling and simulator learning gain.

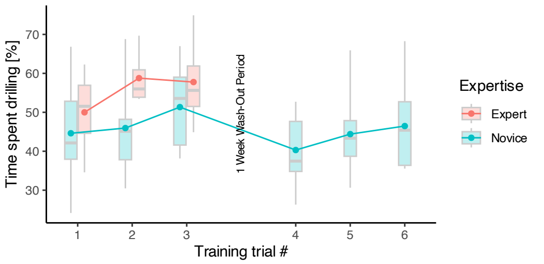

Figure 12: Time spent drilling over time. The participants' time allocated towards drilling (proportionally) over time and grouped by expertise.

familiarization session, we allowed participants to do the procedure once without scoring it, to allow them to get used to the simulator. After each training trial, they could inspect their drilling result in detail on a separate computer screen. The first training session using the simulator, consisting of three trials, took place on day one after the pre-test was conducted and they familiarized themselves. The drilling time for each trial was on average (4.38 ± 2.54) min, but in total with all preparations it took around 5 min to 30 min. Once per participant, we measured the participant's interpupillary distance (IPD) via a depth-sensor (average IPD was (61.64 ± 2.92) mm). Before each trial, we adjusted the HMD IPD slider according to the participants' IPD, calibrated the eye-tracking sensor, and performed an eye-tracking accuracy check. The second training session of three trials with the simulator, along with the follow-up post-test, took place afterward on Day 2, the same day. There was a washout period of (7.53 ± 2.74) d between training days 1 and 2. Two experts independently scored the pre- and post-test plastic teeth.

5 RESULTS

In this section, we will present the analysis of all data, that was gathered during the user study. We will first look at the learning effect that the training inside the simulator produced, measured by real-world assessment. In the next section, we will consider all data that was gathered by the eye tracking, and how it relates to the performance and learning effectiveness.

5.1 Learning Effect

The pre-training errors are in the range $e_1 \in [8, 14.5]$, with average of $\mathrm{M} = 10.42$ and standard deviation $\mathrm{SD} = 1.72$, subsequently denoted as $\mathrm{M} \pm \mathrm{SD}$. The post-training errors are in the range $e_2 \in [0.5, 6]$, with the average 4.08 ± 1.52. The resulting learning gain is -6.48 ± 1.43 (see Fig. 7). A t-test shows the difference between the two means is significant ($p < .001$, $t(22) = 22.42$). We observed a similar improvement in performance when measured inside the simulator (see Fig. 8). The dental novices decreased their error from 2.62 ± 1.33 on trial #1 to 1.22 ± 0.72 on trial #6. The t-test shows a significance of $p < .001$, $t(22) = 4.79$. However, the dental experts did not improve over three trials, going from 1.47 ± 0.79 to 1.38 ± 1.14, with a t-test showing no significance ($p = .45$, $t(3) = 0.14$). Experts produced significantly lower simulator errors (1.43 ± 0.87), compared to the novices on day 1 (2.1 ± 1.21) ($p < .05$, $t(19.4) = -2.33$). However, on day 2, novices produced a more similar level of errors (1.63 ± 0.93), compared to the experts ($p = .24$, $t(15.7) = -0.73$). The analysis of the correlation of learning gains in real-world and corresponding gains during training (as measured by the automated scoring) shows a high positive correlation (Pearson's $R = 0.60$, $p < .01$) (see Fig. 7). This means someone who experienced large improvements inside the simulator also experienced large real-world improvements. We saw no correlation between absolute time to completion (TTC) and performance ($R = 0.08$, $p = .34$). However, experts' TTC ((2.11 ± 0.98) min) was significantly lower than novices' TTC ((4.64 ± 2.30) min) ($p < .001$, $t(23.7) = -7.39$). Although novices improved during training from (4.96 ± 2.64) min on trial #1 to (4.11 ± 2.32) min on trial #6 ($p = .028$, $t(22) = 2.01$), they never reached TTC comparable to experts.

5.2 Eye Tracking

The central task in root-canal access opening is the removal of material. Therefore, we analyzed the vision behavior during active drilling and its impact on performance and learning. First, we looked at the ratio of correct mirror placement to incorrect mirror placement while active drilling was ongoing. This ratio is significantly higher for experts (53.93 %) compared to novices on day 1 (44.01 %, $p < .05$, $t(25.19) = -1.81$), but not on day 2 (48.57 %, $p = .18$, $t(29.98) = -0.91$) (see Fig. 10). This was a result of novices increasing their correct mirror placement ratio from 39.93 % on trial #1 to 54.70 % on trial #6 ($p = .06$, $t(19.71) = -1.67$). Further analysis shows that a large ratio of correct mirror placement during drilling is moderately correlated with low simulator error

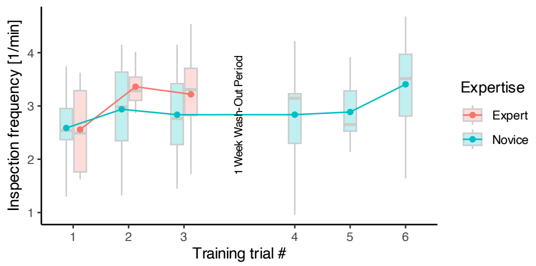

Figure 13: Inspection frequency over time. The frequency of inspections of participants over the training, grouped by expertise.

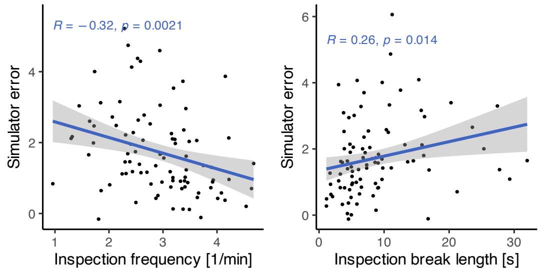

Figure 14: Inspections & simulator error. *Left:* The relation between frequency of inspections and simulator error. *Right:* The relation between average duration of time between inspections and simulator error.

($R = -0.47$, $p < .001$) and increased learning gain ($R = -0.41$, $p < .001$) (see Fig. 9).

We further found that participants' time allocation to drilling and inspection had an impact on the learning effect. In fact, larger allocation towards drilling weakly correlates with decreased learning gain, in real-world learning ($R = 0.22$, $p < .05$) and in simulator learning ($R = 0.34$, $p < .01$) (see Fig. 11). Therefore, a larger allocation of time toward inspection has an negative correlation with increased learning. We did not see a consistent increase or decrease of this metric, going from trial #1 (44.62 %) to trial #6 (46.47 %) ($p = .68$, $t(24.81) = -0.41$) (see Fig. 12). In fact, experts had a significantly larger allocation towards drilling (55.51 %), compared to novices on day 1 (47.36 %) ($p < .05$, $t(19.08) = -2.29$) and on day 2 (43.61 %) ($p < .01$, $t(17.79) = -3.41$).

During the training, we observed most participants were doing regular *inspections* in between periods of drilling. Their frequency and length differed greatly, which made us wonder if it was important for performance. When looking at the frequency of inspections, we observed a learning effect over the time period of the training (see Fig. 13). Novices increased significantly from 2.58 min^{-1} in trial #1 to 3.07 min^{-1} in trial #6 ($p < .05$, $t(36.09) = -2.14$). Experts had a similar improvement from 2.56 min^{-1} in trial #1 to 3.22 min^{-1} in trial #3, although the difference is not statistically significant ($p = .21$). We found a large inspection frequency to moderately correlate with a low simulator error for novices ($R = -0.32$, $p < .01$) (see Fig. 14). On the same note, we found that longer breaks between inspections are weakly correlated with a larger simulator error ($R = 0.26$, $p < .05$). Longer inspection breaks are also weakly correlated with a worse learning gain ($R = 0.21$, $p = .06$). For experts, however, there were no such correlations found.

Next, we looked at the consistency of vision and its impact on performance and learning. A larger amount of significant breaks of vision is associated with worse performance ($R = 0.52$, $p < .001$) and worse learning gain ($R = 0.34$, $p < .01$) in novices (see Fig. 16). Novices improved significantly in this regard, decreasing prolonged vision breaks from 25.35 in trial #1 to 14.26 in trial #6 ($p < .05$, $t(41.55) = 2.08$) (see Fig. 15). Although novices improved significantly, they never reached comparable numbers to the experts. Experts achieved significantly lower prolonged breaks of vision than novices on day 1 ($p < .001$, $t(75.44) = 4.51$) and on day 2 ($p < .001$, $t(66.17) = 4.46$). Experts did not improve significantly in terms of vision breaks during the training session ($p = .34$).

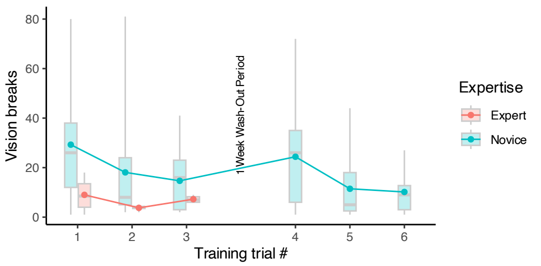

Figure 15: Vision breaks over time. The average amount of prolonged breaks in the vision of participants over the training, grouped by expertise.

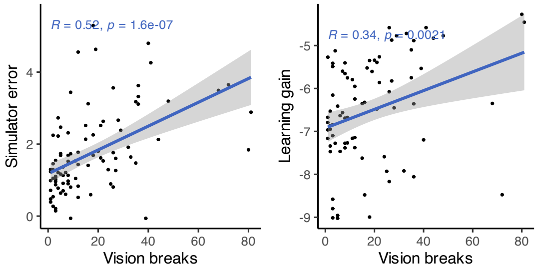

Figure 16: Vision break relations. *Left:* The relation between the amount of prolonged breaks of vision and simulator error. *Right:* The relation between the amount of prolonged breaks of vision and real-world learning gain.

6 DISCUSSION

The learning effectiveness of our simulator is clearly demonstrated by the significant learning gains of participating students. The extent of training transfer between different learning tasks and situations has long been recognized as dependent on task similarity [9]. We created a faithful replica of the patient's mouth and the tools employed in real settings in an effort to ensure a high degree of transfer of psychomotor activities learned in our simulator to those required in surgical procedures. This approach is based on the theory of learning transfer, which posits that repetitive training of a psychomotor activity can yield significant enhancements in the execution of another psychomotor activity with similar components [10,33]. This is further shown by our performance analysis during training, as participants with large training gains tended to also have large real learning gains. This high positive correlation suggests that the learning that happens inside the simulator significantly transfers to the real world. This aligns with other studies that have successfully demonstrated the usefulness of VR simulators for teaching dentistry [6,20,36]. We also evaluated the performance of expert dentists in our simulator to investigate important skill-related behaviors inside the simulator, such as tool usage and gaze behavior. We observed no improvement in the experts' simulator performance, indicating that the skills they acquired over countless real-life sessions transfer well to the simulator. Experts' performance was significantly better than novices on day 1, whereas this difference was statistically insignificant on day 2. This again indicates that our participating students had a significant learning experience over these six simulator trials. We did not reach a learning plateau during the training phase, as especially from the second to last to the last trial, we observed significant performance improvement. However, the novices' performance on the last trial was already comparable to experts performance. Therefore, we are confident, that our training duration of 6 trials is very close to the learning plateau.

Using a VR-based simulator provides the unique ability to track tool usage and gaze behavior with high accuracy and resolution. The placement and use of the mirror during drilling is a cognitive and psychomotor challenge, as the visual occlusion of the handpiece needs to be taken into account, and the handpiece and bur movement are visually mirrored. This leads to limitations of the operator's vision during drilling, as well as high cognivite load. We have found that experts, naturally, had a much higher rate of correct mirror placement and, therefore, a clearer view of the tooth during drilling than novices. This difference was, however, only significant on day one (first three trials) and not on day two, which suggests that the simulator training helped students improve on this metric. Furthermore, we found that a high rate of correct mirror placement during drilling correlates with lower drilling errors in the simulator and increased real-world learning gains. These findings suggest that close attention should be paid to dental students' mirror use during training, and analysis of eye gaze and indirect vision is a useful indicator of skill and learning effectiveness.

Although drilling is the central task in root canal access opening, we have observed that participants spent a considerable amount of time inspecting the tooth (using the mirror) without drilling to get a detailed view of their drilling progress and to plan the next steps in the procedure. Our analysis revealed that spending less time proportionally on drilling, i.e., more time on inspection, correlates well with higher learning gains and training gains. Incidentally, dental experts allocated significantly more time to drilling than students did on day one and day two, reserving less time for inspections. This finding could appear counter-intuitive, but we think it indicates that during the learning phase, spending more time on inspection besides drilling is helpful for learning. The reason for this relation could be that inspections result in timely feedback, which was repeatedly shown to have a positive effect on learning effectiveness [32, 30, 8]. However, spending time on inspection is not important to perform well if one has mastered the procedure. Further, when discretizing inspection events, we saw that the frequency of inspections and breaks between inspections were not significantly different between students and experts, even though both metrics correlated with lower simulator error, at least for students. This finding suggests that the simulator teaches students to inspect frequently. In fact, dental students should be encouraged to frequently inspect the target tooth, but less emphasis is needed once students have mastered the procedure. On a more general level, this could suggest that eye gaze patterns can provide insights into a student's skills and level towards mastery, especially with bi-manual tasks, and even more with indirect vision.

To further analyze the impact of vision, we discretized prolonged breaks in the participant's ability to see either tooth or bur, regardless of the situation. This metric also indicates that the frequency of such vision breaks correlate with more simulator errors and decreased real-world learning gains. Also, students improved significantly w.r.t. this metric over the training period, although this is the only metric in which they never reached levels comparable to the experts. Nevertheless, this finding suggests that continuity of vision is one of the most important factors in a dentist's performance. Understandably, not all intricacies of mirror handling, such as avoiding the occlusion of the handpiece, can be mastered in just six training sessions.

Our findings demonstrate the importance of mastering mirror handling for dentistry students to improve their skills. We also think that our findings suggest that VR surgrey simulators for all kinds of procedures involving indirect vision (e.g., minimally-invasive procedure) could benefit from analyzing students' eye gaze behavior, in order to asssess their learning progress.

In this study we did not give any external feedback, especially while the novices were inside the simulation, operating on the teeth. One could question how in this setting, any improvement of mirror handling can occur. However, when given a task that requires the use of a mirror, repeatedly performing this task facilitates improvement at that task, as shown by [3]. Another problem with external feedback (such as an instructor giving verbal commands, or holding the users hand to guide it) is that its severity and quality might vary greatly, and thus introduce a new source of noise in the generated data. It would be interesting to see the effects of a tutoring system that would automatically compute and display the optimal mirror pose, based on the current viewpoint and the pose of the handpiece. This ideal mirror pose could be rendered visually using kind of a special effect, or using the force feedback device as kind of a guidance (see, for instance, [16]).

Many students were enthusiastic and motivated when using the simulator, because they realized immediately that in the virtual simulator, there is no risk of producing a consequential error. This knowledge can encourage students to practice more, possibly make mistakes, and learn from them, i.e., an otherwise rather stressful situation can be interpreted more as play, which helps to skillfully incorporate their emotions when solving problems [11]. Many students in our study expressed their wish for general access to such a system as part of their training.

7 CONCLUSION

We have presented a novel extension to a VR simulator for complex dental surgeries. The simulator now allows for optical magnification and eye-tracking during bi-manual interaction with realistic haptic feedback. More importantly, we have conducted a user study that shows that faithful VR simulators can be an effective learning tool for acquiring psychomotor bi-manual skills, including indirect vision, as the performance gains inside the simulator translate well into the real world. We have measured real-world learning progress by having students operate on plastic teeth with anatomically correct pulp cavities in a phantom head, which were then rated independently by two dental experts.

Although drilling skills are the primary factor in root-canal access opening, the chosen upper jaw tooth requires extensive and skillful use of a mirror to redirect gaze. Our analysis of eye tracking data reveals that the students' mirror-handling approaches that of the experts with increasing training. In particular, they improved on the frequency of detailed progress inspections and reduced the prolonged periods in which they operated without vision. This suggests that the skillful adjustment of the indirect vision during this kind of bi-manual surgery is a crucial performance measure for learning success. Our findings are novel and provide valuable insights, paving the way for more effective dental and surgical education using VR simulators in the future. This is based on the evidence we found, because it allows the teachers not only to measure *if* the student succeeded, but also *why*. Such insights could provide hints on how to improve training sessions.

This opens up exciting avenues for future work. For instance, it could be possible to enhance our simulator by adding a tutoring system for the mirror and, hence, actively support the learning process. Moreover, we would like to investigate if our findings also hold for other complex asymmetric bi-manual tasks such as laparoscopic surgeries. Finally, it would be interesting to see if other eye gaze metrics can provide more information about a student's learning status, for instance if and where they might struggle with a specific step.

REFERENCES

[1] I. Badash, K. Burtt, C. A. Solorzano, and J. N. Carey. Innovations in surgery simulation: a review of past, current and future techniques. *Annals of translational medicine*, 4(23), 2016. 1

[2] N. Castner, L. Geßler, D. Geisler, F. Hüttig, and E. Kasneci. Towards expert gaze modeling and recognition of a user's attention in realtime. *Procedia Computer Science*, 176, 2020. 2

[3] F. Chu, J. Zheng, Q. Wang, X. Lu, Y. Chen, Y. Zhong, Y. Li, J. Shi, Y. Jiang, W. Zhang, et al. Mirror training device improves dental students' performance on virtual simulation dental training system. *BMC Medical Education*, 23(1):315, 2023. 2, 9

[4] F. Costa, J. Pacheco-Yanes, J. Siqueira Jr, A. Oliveira, I. Gazzaneo, C. Amorim, P. Santos, and F. Alves. Association between missed canals and apical periodontitis. *International Endodontic Journal*, 52(4):400–406, 2019. 2

[5] M. J. Díaz, E. Sánchez, J. J. Hidalgo, J. M. Vega, and M. Yanguas. Assessment of a preclinical training system with indirect vision for dental education. *European Journal of Dental Education*, 5(3):120–126, 2001. 2

[6] E. J. Eve, S. Koo, A. A. Alshihri, J. Cormier, M. Kozhenikov, R. B. Donoff, and N. Y. Karimbux. Performance of dental students versus prosthodontics residents on a 3D immersive haptic simulator. *Journal of dental education*, 78(4):630–637, 2014. Publisher: Wiley Online Library. 1, 8

[7] F. Freyberger, M. Kuschel, B. Färber, and M. Buss. A presence measure for virtual reality and telepresence based on multimodal conflicts. In *Proceedings of the ISPR Presence 2007. The 10th Annual International Workshop on Presence*, Oct. 2007. 3

[8] E. R. Fyfe and B. Rittle-Johnson. Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology*, 108(1):82, 2016. 9

[9] R. M. Gagne, K. E. Baker, and H. Foster. On the relation between similarity and transfer of training in the learning of discriminative motor tasks. *Psychological Review*, 57(2):67, 1950. 8

[10] R. M. Gagne, H. Foster, and M. E. Crowley. The measurement of transfer of training. *Psychological Bulletin*, 45(2):97, 1948. 8

[11] W. W. Y. Ho. Influence of play on positive psychological development in emerging adulthood: A serial mediation model. *Frontiers in Psychology*, 13, 2022. doi: 10.3389/fpsyg.2022.1057557 9

[12] B. Hosp, Y. Sauer, R. Agarwala, and S. Wahl. Simulation tool for optical systems: Evaluating and optimizing autofocals and beyond. Sept. 2023. doi: 10.21203/rs.3.rs-3318024/v1 2

[13] B. Hosp, M. S. Yin, P. Haddawy, R. Watcharopas, P. Sa-ngasoongsong, and E. Kasneci. Differentiating surgeons' expertise solely by eye movement features. In *Companion Publication of the 2021 International Conference on Multimodal Interaction*, ICMI '21 Companion, p. 371–375. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3461615.3485437 2, 5

[14] M. Jeong, B. Kim, and J. Ryu. The effects of expertise level on performance time, task accuracy, and task load in virtual reality dental simulation. *The Korean Journal of Educational Methodology Studies*, 32-2:185–204, 2020. 2

[15] M. Kaluschke, R. Weller, N. Hammer, L. Pelliccia, M. Lorenz, and G. Zachmann. Realistic haptic feedback for material removal in medical simulations. In *2020 IEEE haptics symposium (HAPTICS)*, pp. 920–926, 2020. doi: 10.1109/HAPTICS45997.2020.ras.HAP20.74.13165668 3

[16] M. Kaluschke, M. S. Yin, P. Haddawy, N. Srimaneekarn, P. Saikaew, and G. Zachmann. A Shared Haptic Virtual Environment for Dental Surgical Skill Training. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 347–352. IEEE, Lisbon, Portugal, Mar. 2021. doi: 10.1109/VRW52623.2021.00069 9

[17] M. Kaluschke, M. S. Yin, P. Haddawy, S. Suebnukarn, and G. Zachmann. The Impact of 3D Stereopsis and Hand-Tool Alignment on Effectiveness of a VR-based Simulator for Dental Training. In *2022 IEEE 10th International Conference on Healthcare Informatics (ICHI)*, pp. 449–455. IEEE, Rochester, MN, USA, June 2022. doi: 10.1109/ICHI54592.2022.00067 2

[18] M. Kaluschke, M. S. Yin, P. Haddawy, S. Suebnukarn, and G. Zachmann. The effect of 3d stereopsis and hand-tool alignment on learning effectiveness and skill transfer of a vr-based simulator for dental training. *PLOS ONE*, 18(10):1–25, 10 2023. doi: 10.1371/journal.pone.0291389 6

[19] G. W. Knight and P. J. Guenzel. Design and validation of mirror skills instruction. *Journal of Dental Education*, 58(10):752–761, 1994. 2

[20] P. Koopman, J. Buis, P. Wesselink, and M. Vervoorn. Simodont, a haptic dental training simulator combined with courseware. *Bio Algorithms Med Syst.*, 6(11):117–122, 2010. 1, 8

[21] R. S. Kunovich and R. G. Rashid. Mirror training in three dimensions for dental students. *Perceptual and motor skills*, 75(3):923–928, 1992. 2

[22] A. Leube, K. Rifai, and K. Rifai. Sampling rate influences saccade detection in mobile eye tracking of a reading task. *Journal of Eye Movement Research*, 10(3):10.16910/jemr.10.3.3, 2017. doi: 10.16910/jemr.10.3.3 2, 4

[23] D. Lugassy, Y. Levanon, R. Pilo, A. Shelly, G. Rosen, A. Meirowitz, and T. Brosh. Predicting the clinical performance of dental students with a manual dexterity test. *PloS one*, 13(3):e0193980, 2018. 2

[24] A. R. McClure, T. C. Roomian, S. E. Eisen, G. Kugel, and R. B. Amato. Jumpstart mirror trainer: a new device for teaching mirror skills to first-year dental students. *Journal of Dental Education*, 83(10):1199–1204, 2019. 2

[25] M. Meraner and J. B. Nase. Magnification in dental practice and education: experience and attitudes of a dental school faculty. *Journal of Dental Education*, 72(6):698–706, 2008. 2

[26] L. Neumann. A simple exercise for teaching mirror vision skills. *J Dent Educ*, 52:170–172, 1988. 2

[27] P. Perrin, K. Neuhaus, and A. Lussi. The impact of loupes and microscopes on vision in endodontics. *International endodontic journal*, 47(5):425–429, 2014. 2

[28] G. M. Rau and A. K. Rau. Training device for dental students to practice mirror-inverted movements. *Journal of Dental Education*, 75(9):1280–1284, 2011. 2

[29] I. Schuetz and K. Fiehler. Eye Tracking in Virtual Reality: Vive Pro Eye Spatial Accuracy, Precision, and Calibration Reliability. *Journal of Eye Movement Research*, 15(3):10.16910/jemr.15.3.3, 2022. doi: 10.16910/jemr.15.3.3 5

[30] V. J. Shute. Focus on formative feedback. *Review of educational research*, 78(1):153–189, 2008. 9

[31] A. Sipatchin, S. Wahl, and K. Rifai. Eye-Tracking for Clinical Ophthalmology with Virtual Reality (VR): A Case Study of the HTC Vive Pro Eye's Usability. *Healthcare*, 9(2):180, Feb. 2021. doi: 10.3390/healthcare9020180 5

[32] S. Taxipulati and H.-D. Lu. The influence of feedback content and feedback time on multimedia learning achievement of college students and its mechanism. *Frontiers in Psychology*, 12, 2021. doi: 10.3389/fpsyg.2021.706821 9

[33] E. L. Thorndike. The effect of practice in the case of a purely intellectual function. *The American Journal of Psychology*, 19(3):374–384, 1908. 8

[34] E. Yang, S. Park, J. Ryu, and T. Lim. How does dental students' expertise influence their clinical performance and perceived task load in a virtual dental lab? *Journal of Computing in Higher Education*, pp. 1–27, 2022. 2, 5

[35] M. S. Yin, P. Haddawy, B. Hosp, P. Sa-ngasoongsong, T. Tanprathumwong, M. Sayo, S. Yangyuenpradorn, and A. Supratak. A study of expert/novice perception in arthroscopic shoulder surgery. In *Proceedings of the 4th International Conference on Medical and Health Informatics*, ICMHI '20, p. 71–77. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3418094.3418135 2, 5

[36] M. S. Yin, P. Haddawy, S. Suebnukarn, F. Kulapichitr, P. Rhienmora, V. Jatuwat, and N. Uthaipattanacheep. Formative feedback generation in a VR-based dental surgical skill training simulator. *Journal of Biomedical Informatics*, 114:103659, 2021. Publisher: Elsevier. 1, 8