The Impact of 3D Stereopsis and Hand-Tool Alignment on Effectiveness of a VR-based Simulator for Dental Training

Maximilian Kaluschke
Computer Graphics and Virtual Reality
University of Bremen
Bremen, Germany
mxkl@cs.uni-bremen.de

Myat Su Yin
Faculty of ICT
Mahidol University
Bangkok, Thailand
myatsu.yin@mahidol.ac.th

Peter Haddawy
Faculty of ICT
Mahidol University
Bangkok, Thailand
peter.had@mahidol.ac.th

Siriwan Suebnukarn
Faculty of Dentistry
Thammasat University
Bangkok, Thailand
ssiriwan@tu.ac.th

Gabriel Zachmann
Computer Graphics and Virtual Reality
University of Bremen
Bremen, Germany
zach@cs.uni-bremen.de

***Abstract*—Recent years have seen the proliferation of VR-based dental simulators using a wide variety of different VR configurations. Differences in technologies and setups used result in important differences in degree of realism. These include 3D stereoscopic rendering and visual alignment of the user's hands with the virtual tools. While each new dental simulator typically is associated with some form of evaluation study, only few comparative studies have been carried out to determine the benefits of various simulation aspects.**

In this paper, we seek to determine the impact of 3D stereoscopic rendering and of hand-tool alignment on the teaching effectiveness of a VR dental simulator. We developed a bimanual simulator using an HMD and two haptic devices that provides an immersive environment with both 3D stereoscopic rendering and hand-tool alignment. We then systematically and independently controlled for each of the two aspects of the simulation. We trained four groups of students in root canal access opening using the simulator and measured the learning gains by doing pre- and post-testing using realistic plastic teeth. We found that hand-tool alignment has a positive impact on learning gains, while stereoscopic 3D does not. The effect of stereoscopic 3D is surprising and demands further research in settings with small target objects. The results of our study provide valuable information for the future design of dental simulators, as well as simulators for other high-precision psycho-motor tasks.

***Index Terms*—virtual reality, dental training, stereoscopic rendering, hand-tool alignment**

I. Introduction

Dental schools have long used simulation to provide students with deliberate practice of particular skills before practicing on live patients. The early, and still most commonly used, simulators consist of a mannequin head (called a phantom head) fitted with plastic teeth. Students can then use real dental instruments (drill, mirror) to practice relatively simple procedures such as caries removal using inexpensive simple solid plastic teeth and more complex procedures such as root canal access opening using more expensive plastic teeth with internal anatomy.

Recent years have seen the proliferation of VR-based dental simulators due to enabling technological advancements, combined with concrete benefits of the approach [1], [2], [3]. VR simulators offer high-fidelity simulations that are reusable and can be configured to provide trainees practice on a variety of different cases. They also have the ability to record accurate data on individual performance, which provides the opportunity for trainees to receive objective feedback. With the variety of VR technologies available, dental simulators have been developed using a wide variety of different VR configurations. Display technologies used include traditional 2D monitors, 3D monitors, half-mirrored displays, and head-mounted displays (HMD), the latter three of which provide stereoscopic depth perception. Instrument manipulation is achieved with and without haptic feedback. In addition, the use of HMDs and half-mirrored displays supports hand-tool alignment in which the user sees the dental instrument in the same location as their physical hand. In contrast, 2D and 3D monitors do not provide such alignment. While each new dental simulator typically is associated with some form of evaluation study [4], [5], [6], only few comparative studies have been carried out to determine the benefits of the simulation aspects associated with the various available VR technologies being used and none that examine the impact on transferability of learned skills.

In this paper we seek to determine the impact of 3D stereoscopic rendering and of hand-tool alignment on the teaching effectiveness of a VR dental simulator. We developed a simulator using an HMD and two haptic devices that provides an immersive environment with both 3D stereoscopic rendering and hand-tool alignment. We then systematically and independently controlled for each of the two aspects of the simulation. We trained four groups of students in root canal access opening over a period of two days using the simulator and measured the learning gains by doing pre- and post-testing using realistic plastic teeth. We find, not surprisingly, that hand-tool alignment has a impact on learning gains, while, somewhat surprisingly, that stereoscopic 3D does not. We hypothesize that the result concerning 3D stereoscopic rendering may be related to the resolution of the HMD. The results of our study provide valuable information for the future design of dental simulators, as well as simulators for other high-precision psycho-motor tasks.

II. Related Work

With an increasing trend of using 3D stereo-projected images to create realistic virtual learning environments, there is an ongoing debate as to whether stereo-projected images are a necessary feature of simulators [7], [8], [9]. A comprehensive review conducted by McIntire et al. [7], found that in 15% of over 180 experiments from 160 publications, stereoscopic 3D display either showed a marginal benefit over a 2D display or the results were mixed or unclear, while in 25% of experiments, stereoscopic 3D display showed no benefit over non-stereo 2D viewing. They concluded that stereoscopic 3D displays are most useful for tasks involving the manipulation of objects and for finding/identifying/classifying objects or imagery. de Boer et al. [10] investigate the differences in students' performance in carrying out manual dexterity exercises with the Simodont dental trainer simulator (The MOOG Industrial Group; www.moog.com) in 2D and 3D versions. 3D vision in the dental trainer was based on the projection of two images superimposed onto the same screen through a polarising filter. 2D vision was obtained by turning off one of the two projectors such that only one image was projected onto the screen. All of the students in both the 2D and 3D vision groups wore polarised glasses during the practice sessions and when testing to keep the environmental factors constant. The task consisted of using a dental drill to remove material from a cube and the outcomes were automatically scored. The results showed that students working with 3D vision achieved significantly better results than students who worked in 2D. In an administered questionnaire, participants also indicated that they preferred the 3D vision setting. Students reported having an unpleasant experience in working with 2D vision while wearing the glasses. The probable reason is that only one eye received an image through the polarized glasses. In a related study, Al-Saud and colleagues[11] examined the effects of stereopsis on dentists' performance with the Simodont dental simulator. Thirteen qualified dentists were recruited and asked to performed a total of four different dental manual dexterity tasks under non-stereoscopic and stereoscopic vision conditions with direct and indirect (mirror) observation. The tasks consisted of removal of material from a geometric shape embedded in a cube of material. Automated scoring was based on amounts of target and non-target material removed. Stereoscopic 3D was the simulator's normal operation and was achieved as in the previously mentioned study in [10]. To produce 2D images, the simulator was engineered to output a single image to both eyes. The study found out that depth related errors were significantly higher under non-stereoscopic viewing but lateral errors did not differ between conditions.

Collaco et. al. [12] investigated the effects of (full) immersion and haptic feedback on inferior alveolar nerve anesthesia technical skills training. Their experimental study consists of preceptorship and training phases. During the preceptorship phase, one of the groups received the anesthesia instructions from the dental instructor on a full HD TV screen, while the participants from the remaining three groups observed the anesthesia procedure from the instructor's perspective in immersive condition using the HMD. In the training phase, the participants in one of the groups in immersive condition during the previous preceptorship stage performed the anesthesia injection using the full HD TV screen while the remaining three groups performed the task with the HMD in immersive condition. Their findings include that the participants from immersed groups either in preceptorship or training performed the anesthetic procedure faster and more accurately than those in the combined non-immersed groups. The authors conclude that the stereoscopic view from the HMD in immersed groups provided a better perception of depth when compared to the 2D monitor, making instrument navigation inside the mouth easier and leading to better performance results.

In manipulating tools, users receive information from two feedback loops: the body-related proximal feedback loop (proximal action effect) such as tactile sensations from the moving hand, and from the effect in distal space, such as the visual feedback from the movement of effect points of the tool (distal action effect). Establishing the mapping between the moving hand and the moving effect part of the tool can add challenges to the human information processing systems. According to Sutter et al. [13], if information from proximal and distal feedback loops are equally important for controlling actions, any discrepancy between them would be a constant source of interference to the user. Users of conventional desktop-based VR simulators using haptic interfaces are familiar with this scenario while manipulating the haptic device and observing the action effects on a display monitor. Meanwhile, in HMD simulators, the spatial gap between the hand and the resulting movement can be eliminated by manipulating the virtual camera position to the user in such a way that the user sees and feels as if he is manipulating the dental tools on the patient's teeth. Although more realistic, it is interesting to note that in this condition the vision may be afforded with a higher weighting than other sensory information; a situation often referred to as visual capture. Although visual information is invaluable for executing skillful manual tasks, visual capture can produce powerful illusory effects with individuals misjudging the size and position of their hands. Moreover, if vision of the hand/tool is available in the operating area it should be recognized that there might well be interference that would impair motor performance and learning, as there is a shift in attentional focus to the outcome of actions rather than the actions themselves. Wilke et al. [14] studied whether

Fig. 1. The VR dental surgery simulator is used by a dentistry student to practice root canal access opening on tooth #36. The VR HMD and haptic input/output devices allow for a intuitive control with realistic haptic feedback (in alignment condition). The monitor shows the image that the student is seeing on the HMD.

Fig. 2. The calibration of the haptic devices with the HTC Vive VR system is implemented using a VR controller with a static offset. The ”hand-tool misalignment” is achieved by calibrating and then moving the haptic devices forward and downward in front of the table, as shown here.

visual capture can interfere with an individual’s rate of motor learning in a laparoscopic surgery setting. They investigated the adaptation to distorted visual feedback in two groups: a direct group directly viewed the input device, while the indirect group used the same input device but viewed their movements on a remote screen. When distortion exists between hand and tool movement, then visual capture is an issue and participants in the indirect group performed better than those in the direct group. However, when no distortions were applied, participants in the direct group performed better than participants in the indirect group. In the dental domain, there is typically no distortion present for drilling tasks. Similarly, Sutter et al. [13] conducted several experiments aiming to investigate the underlying motor and cognitive processes and the limitations of visual predominance in tool actions. Their major finding is that when transformations are in effect the awareness of one’s own actions is quite low.

III. Simulator

We developed a VR dental surgery simulator with haptic feedback, in which students can practice caries removal, crown preparation, and root canal access opening (see Fig. 1). The simulator was developed using Unreal Engine (UE) 4.26. An HTC Vive Pro Eye with a combined resolution of 2880×1600 was used to display stereo images from the UE SteamVR 1.0 plugin. The dental virtual hand-piece and mirror are each controlled by a Geo-Magic Touch haptic device (Phantom) with 6 degrees-of-freedom (DOF) input and 3 DOF output. Haptic feedback is provided to simulate the interaction between the hand piece and virtual tooth. The sound of the drill is also simulated. The virtual patient was modeled using the Metahuman framework [15] and imported into our UE scene. The virtual human is rendered with high fidelity visuals including subtle idle animations of the face and mouth, such as eye blinking and movement of the tongue. We made sure to not include animations that would alter the location of the tooth. We added a transparency texture to the virtual teeth texture, which allows us to hide one of the teeth (#36) of the Metahuman model. In its place, we inserted a new tooth that we modeled by hand with guidance from CT scans of similar teeth and approved by an expert dentist. At runtime, we render the tooth by using the UE Procedural Mesh Component 1.0 (PMC). We generate triangles of modified tooth regions in a CUDA library, which are then fed to UE’s PMC. The library approximates the tooth surface by a metaball surface that is discretized at runtime using a parallel marching cubes implementation with a resolution of $90 \times 135 \times 90$.

We compute the haptic feedback outside of the UE main loop, so as not to be limited by the rendering frame rate. The force is computed according to the algorithm presented in [16], which uses an inner spheres volume representation. The tooth enamel is made up of 100k, the dentin by 170k, and the pulp by 10k spheres. We tuned the force, drilling, and friction parameters by our subjective impression of drilling the real plastic teeth that students usually practice on, with approval by an expert dentist.

A. Stereo Rendering

The standard VR rendering is set up to be at a realistic scale, such that the user has a natural stereo impression from the two different images that are sent to the eye. This setting will later be referenced as the ”stereo” condition during the user study. To investigate the effect that stereo vision has on the learning effect, we implemented a rendering mode that renders the virtual scene without stereoscopy. Our first idea to implement this was to decrease the VR projection inter-pupillary distance (IPD) to 0, however this parameter can not be modified in UE in VR. Another possibility would be to have a screen-space shader that blanks one eye. However, we found, similarly to [10], that it creates an unpleasant feeling. Therefore, we decided to change the effective IPD by modifying the world-to-meters scale in UE from the default 100 to 0.01, thereby increasing the virtual scale by a factor of 10000. This requires to compensate for this new scale to make the player movement and general impression feel normally dimensioned. We achieve this by scaling the VR

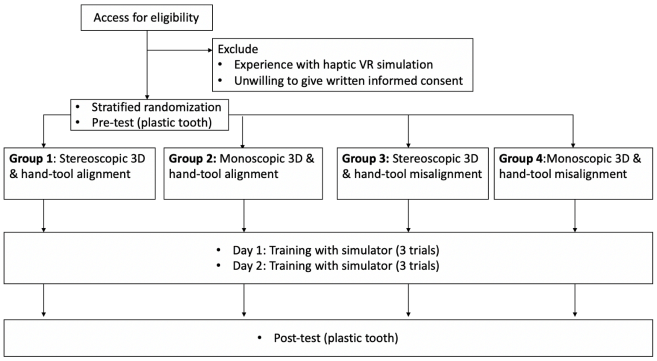

Fig. 3. Flowchart that shows the user study procedure.

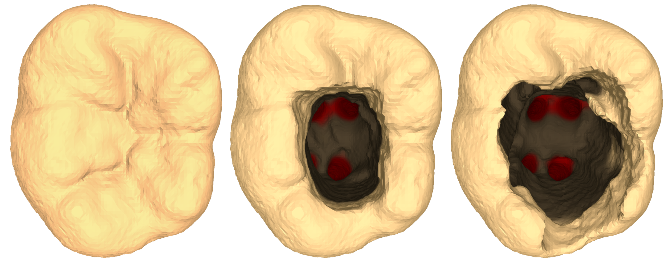

Fig. 4. Different conditions of tooth #36. *Left*: Intact tooth crown. *Center*: A root canal access opening with a low error score. All orifices are accessible with little over-drilling. *Right*: A root canal access opening with a high error score, as multiple walls are over-drilled and not smooth.

HMD and VR controller translational movements by factor 10000, too. The absolute IPD for the renderer however remains unchanged. Thus, we get an effectively very small IPD in a seemingly normally scaled virtual scene. This setting will later be referenced as the "mono" condition during the user study.

B. Hand-Tool Alignment

The force feedback devices are registered with the HTC Vive VR system by using a VR controller dock that is mounted on a board with a static offset to both haptic device bases (Fig. 2 shows the misalignment condition). Inside the game engine, we define the virtual position of the haptic device bases inside the scene. When we run the simulator in a new VR configuration (new light house poses or new haptic device locations), the difference between virtual and real haptic devices is added to the VR camera. By doing this, we overlay the visual tool and haptic device handle from the view of the user. We call this condition "hand-tool alignment" (as shown in Fig. 1). To define the contrasting condition, "hand-tool misalignment", we move the haptic devices after doing the calibration. We moved the haptic devices down by 20 cm and forward by 50 cm (see Fig. 2). We chose this offset to simulate a misalignment setting that resembles the offset on a desktop monitor in VR.

IV. User Study

After receiving ethical approval from the Institutional Review Board from Mahidol and Thammasat universities, we invited students enrolled in the Faculty of Dentistry of Thammasat University to participate in our study. We recruited 40 participants (12 male, 28 female) and conducted a randomized controlled study. All participants were fifth year dental students. They were not admitted to the study if any of the following criteria were presented: (i) had received prior experience with the simulation, or (ii) received below 70% marks in knowledge assessment of the endodontic cavity preparation. The participants were randomly assigned to one of the four groups: Group 1: Stereoscopic 3D & hand-tool alignment, Group 2: Monoscopic 3D & hand-tool alignment, Group 3: Stereoscopic 3D & hand-tool misalignment, and Group 4: Monoscopic 3D & hand-tool misalignment. The task for the participants was to perform access opening on the virtual tooth during the training session and on a plastic tooth (lower left molar; tooth number 36; http://www.nissin-dental.net/) in pre- and post-training assessment sessions. The plastic teeth closely resemble the feeling of drilling real teeth and are anatomically correct. A students ability to perform the root canal access opening on such plastic teeth will predict with high reliability his or her ability to perform the task on real human tooth. Thus, using plastic teeth are the best option to assess real-world dental skills that is also ethically sound. Participants were briefly instructed on the use of the simulator, the experiment flow and the requirements of the access opening. As shown in the study flowchart (Fig. 3), the training of each participant took place on two separate days. The first training session using the simulator, consisting of three trials, took place on Day 1, after the pre-test was conducted. Each trial took around 5-30 minutes. The second training session of three trials with the simulator, along with the follow-up post-test, took place afterwards on Day 2, the same day. There was a gap of four to seven days between days 1 and 2 of training. The pre- and post-test plastic teeth were independently scored by two experts. Each of the four cardinal tooth walls and the pulp floor was visually observed and scored for errors by each expert. The criteria for scoring the errors can be summed up in the following way:

- +0 Access to all orifices without an excess cavity.
- +1 Access to all orifices with minor over-drilling.
- +2 Incomplete removal of pulp chamber roof and/or excessive over-drilling.
- +3 Unidentified canals and/ or perforation.

The overall score for a tooth is taken to be the sum of the error scores of the walls and pulp floor. Therefore, the score ranges from 0 to 15, with lower values indicating better performance (examples shown in Fig. 4). The experts' scores had excellent reliability ($\kappa = 0.87$, and intra-class correlation of 0.98). Therefore, we used the mean value of the two experts' scores in the analysis.

V. Results

The error score for the pre-test ranges from 1 to 6.5, whereas the post-test scores ranges from 0 to 7 (see Fig. 5). We have two hypotheses when looking at the gathered data: Stereo

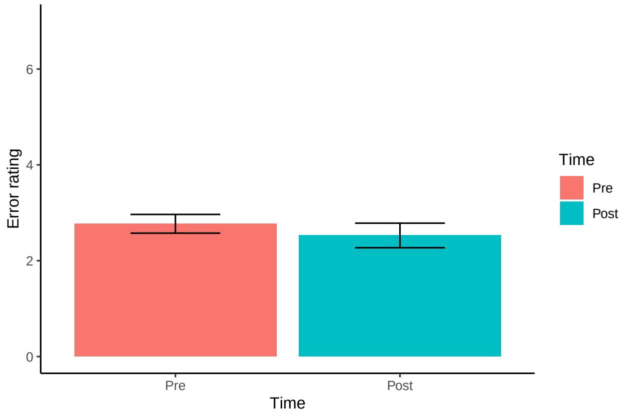

Fig. 5. The paired error scores of all students before and after training.

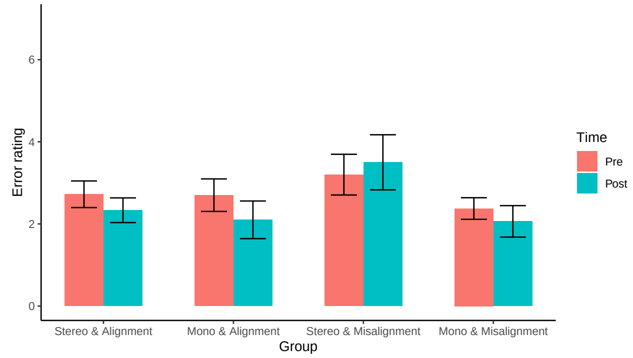

Fig. 6. The paired error scores of all students before and after training.

rendering improves the learning gain, and hand-tool alignment improves the learning gain. We first define the error change $e_\triangle$ for each student as the difference between pre-test error score, e_0, and post-test error score, e_1, so $e_\triangle = e_1 - e_0$. With this, $e_\triangle$ defines the inverse learning gain for each student. The learning gain is normally distributed around $M = -0.375$ with $SD = 1.87$, ranging between -5 and 4. We determined 3 outliers based on interquartile range analysis, resulting in removing the following learning gains: $\{-5, -5, +4\}$. Even though these outliers improve the apparent effectiveness of our dental surgery simulator, they are very unusually high and low learning gains which we feel do not represent an effect of the participant group but rather an inherent property of the participant. After removing outliers, the distribution is centered around the slightly larger $M = -0.24$ with standard deviation $SD = 1.43$.

Looking at the pre- and post-scores, we observe a small overall decrease of students' error score from pre ($M = 2.77$, $SD = 1.19$) to post ($M = 2.53$, $SD = 1.56$) root canal access opening. A paired one-tailed t-test shows a mean difference of -0.2432, with significance of $p = 0.153$. Based on the p-value, we can not determine whether the students' overall improvement in performance is caused by the training.

A. Groups

Between the four groups (as detailed in IV) we found differences in how much participants learned the task of root canal access opening. To determine the learning effect we compare each participants' pre-test error score to their post-test error score. The statistical significance is determined here by a paired one-tailed t-test with the hypothesis that the post-test error scores are lower than the paired pre-test error scores. As the learning gain is normally distributed, we used the parametric t-test. The means and standard deviations of pre- and post-test error rating per group are visualized in Fig. 6. The significant tests showed that none of the learning effects of the four groups are statistically significant.

We found that participants of group "stereo & alignment" performed slightly better at the post-test ($M = 2.33$, $SD = 0.901$) compared to the pre-test ($M = 2.72$, $SD = 0.972$). The difference of the mean error scores is -0.389. Participants of group "mono & alignment" improved their drilling performance between pre ($M = 2.7$, $SD = 1.25$) and post test ($M = 2.1$, $SD = 1.45$). The difference in error score of -0.6 is substantial. Participants of group "stereo & misalignment" on average scored worse in the post-test ($M = 3.5$, $SD = 2.12$) compared to the pre-test ($M = 3.2$, $SD = 1.57$). The difference in error score is $+0.3$. The scores of participants in group "mono & misalignment" have improved in the post-test ($M = 2.06$, $SD = 1.08$) compared to the pre-test ($M = 2.38$, $SD = 0.744$). This is an improvement of -0.312 in the error score. A one-way ANOVA showed no statistically significant differences between the mean learning gains of the groups ($F(3, 36) = 1.436$, $p = 0.552$).

B. 3D Rendering Modes

We also looked at the effect that stereoscopic rendering had on the participants' performance (see Fig. 7). Here, we regard the data of group 1 & 3 as one set of data ("stereo"), and 2 & 4 as the other set of data ("mono"). We thereby control for the alignment condition. The "stereo" group's pre-test error scores ($M = 2.97$, $SD = 1.31$) decreased by 0.0263

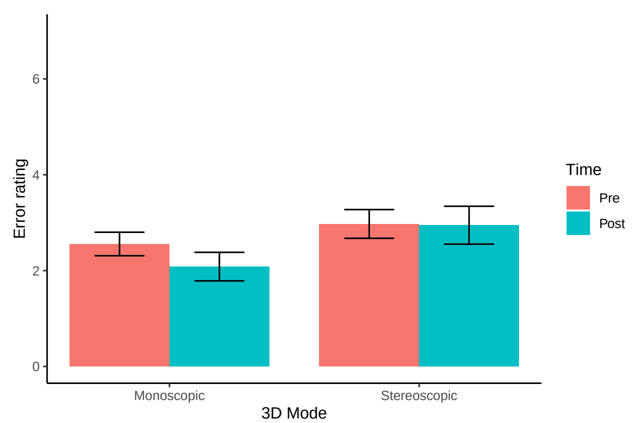

Fig. 7. The paired error scores of all students before and after training, grouped by the stereo factor.

for the post-test ($M = 2.95$, $SD = 1.72$). The one-tailed t-test showed that the increase is likely a result of random chance ($p = 0.4695$). Therefore the students in the "stereo" group did not improve because of the training. In contrast, the "mono" group's post-test error scores ($M = 2.08$, $SD = 1.26$) improved compared to the pre-test error scores ($M = 2.56$, $SD = 1.04$). This large difference of -0.472 have a statistical significance of $p = 0.082$. This means the students of the "mono" group did improve because of the training in VR. This suggests that students performed better after training in the "mono" condition, which is not the case for the "stereo" condition. To measure the effect of the 3D rendering mode on the learning effectiveness we compared the mean learning gains using a parametric two-tailed t-test. The differences of means of the learning gain between "mono" ($M = -0.472$) and "stereo" ($M = -0.026$) is 0.446, however the difference is not statistically significant ($p = 0.349$).

C. Hand-Tool Alignment

The other factor we looked at is the hand-tool alignment and what impact it had on the participants' performance (see Fig. 8). Here, we regard the data of group 1 & 2 as one set of data ("aligned"), and 3 & 4 as the other set of data ("misaligned"), controlling for the stereo factor. The misalignment group did slightly worse on their post-test ($M = 2.86$, $SD = 1.85$), compared to their pre-test ($M = 2.83$, $SD = 1.31$). This small difference of 0.0278 was however shown by the t-test to be likely by random chance ($p = 0.5307$). Therefore the participants of the group "misalignment" did not improve by virtual training. However, the "alignment" group improved from their pre-test ($M = 2.71$, $SD = 1.31$) by -0.5 for their post-test ($M = 2.21$, $SD = 1.19$). The t-test shows a statistical significance of $p = 0.06$. This suggests that the participants of the "alignment" group improved their score because of the virtual drilling training. This shows, that virtual hand-tool alignment is important for effective training using a virtual simulator.

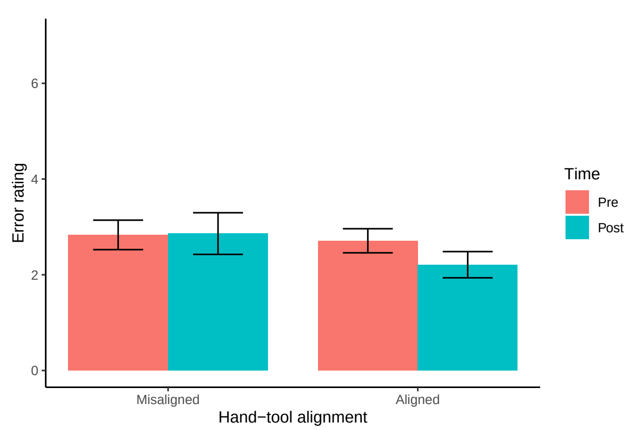

Fig. 8. The paired error scores of all students before and after training, grouped by the hand-tool alignment factor.

VI. Discussion

We observed a considerable difference in performance for the factor "stereo". Our hypothesis was, that people would learn better in the more realistic setting with stereoscopic 3D rendering, since stereoscopic 3D provides the user with additional depth cues. Surprisingly, the group that had monoscopic 3D performed better. Several participants complained about the image being perceived as blurry. However, we don't have objective data if this was more the case for participants that had stereoscopic 3D. [17] showed that there is an upper disparity limit until which binocular vision works effectively. We suspect that this upper limit could be even lower in an HMD like the HTC Vive Pro Eye. In that case, students needed to have a certain minimal distance to the tooth surface in order to limit the inter-ocular disparity. This in turn means that the projected screen size of the tooth has an upper bound, which, given the limited resolution, results in the perception of a blurry image. The monoscopic 3D implementation does not limit the focus range at all, since the eyes do not need to accommodate. Therefore, the projected screen size of the tooth can be freely chosen by moving closer or further to the tooth. This gives an advantageous effective resolution of the tooth for monoscopic 3D when the participants move their head closer to the tooth. This observation indicates that stereoscopic 3D shown on an HMD might not be suitable for detail orientated tasks like surgery, when it involves small objects that need to be seen very closely. In the future, we plan to investigate the effect of HMD resolution on user performance on training intricate manual skills. We hypothesize that an HMD with high resolution will alleviate most of this issue. A zoom feature might also be a useful addition and could be naturally motivated by dentists' tendency to use binocular loupes.

The hand-tool alignment had a similarly large effect size. The students that practiced with aligned virtual and physical tool performed much better. As the misalignment group was designed to be similar to traditional monitor-based simulators, our finding indicates that VR simulators with proper alignment of the virtual and physical world has a major advantage for the students' ability to learn, compared to traditional monitor-based simulators. This finding is consistent with the finding of Wilke et al. [14] in the context of laparoscopic surgery.

A number of factors distinguish this study from previous ones. Previous studies [10], [11] examined performance and learning differences in dental simulators with stereoscopic and monoscopic rendering. In those studies the task was carried out on simulated geometric objects. Evaluation of skill was done within the simulator, with automated scoring based on material removed. In contrast, our study used the endodontic task of root canal access opening. Evaluation of learning gains was done using pre- and post-testing on realistic plastic teeth, with scoring done by dental instructors using the standard method used in clinical teaching. Thus, it can be argued that our study is done in a more realistic setting and includes evaluation of transferability of learned skills. Transferability is important to evaluate since it is fully possible to attain a high level of skill in a simulator, yet not in realistic settings.

VII. Conclusion

We have shown that a VR simulator could be a good tool to help teach students how to acquire dental surgical skills such as root canal access opening. This is the first study to analyse the transferability of dental skills from virtual VR simulation training to real-world learning gains. We have found that the alignment of the physical and virtual tools had a positive impact on students' learning gains, compared to students with misaligned physical and virtual tools. Surprisingly, we observed that in our setting, monoscopic 3D rendering offered students with more helpful training compared to stereoscopic 3D, as their learning gain was higher. We can not generalize these findings to the broader population, as differences are not statistically significant. In our setting, the tooth is a rather small target object, as such the limited HMD resolution could be cofounder of our results, since in other settings, stereoscopic 3D has been shown to improve performance. Further research is needed to understand the effects of stereoscopic 3D in settings with small target objects.

Acknowledgment

This work was partially supported by a grant from the Mahidol University Office of International Relations to Haddawy in support of the Mahidol-Bremen Medical Informatics Research Unit (MIRU), by a fellowship from the Hanse-Wissenschaftskolleg Institute for Advanced Study, and by a Young Researcher grant from Mahidol University to Su Yin.

References

[1] D. Wang, T. Li, Y. Zhang, and J. Hou, "Survey on multisensory feedback virtual reality dental training systems," *European Journal of Dental Education*, vol. 20, no. 4, pp. 248–260, 2016.

[2] Y. Li, H. Ye, F. Ye, Y. Liu, L. Lv, P. Zhang, X. Zhang, and Y. Zhou, "The current situation and future prospects of simulators in dental education," *Journal of Medical Internet Research*, vol. 23, no. 4, p. e23635, 2021.

[3] A. Towers, J. Field, C. Stokes, S. Maddock, and N. Martin, "A scoping review of the use and application of virtual reality in pre-clinical dental education," *British dental journal*, vol. 226, no. 5, pp. 358–366, 2019.

[4] 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*, vol. 78, no. 4, pp. 630–637, 2014.

[5] D. Wang, S. Zhao, T. Li, Y. Zhang, and X. Wang, "Preliminary evaluation of a virtual reality dental simulation system on drilling operation," *Bio-medical materials and engineering*, vol. 26, no. s1, pp. S747–S756, 2015.

[6] 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*, vol. 114, p. 103659, 2021.

[7] J. P. McIntire, P. R. Havig, and E. E. Geiselman, "Stereoscopic 3d displays and human performance: A comprehensive review," *Displays*, vol. 35, no. 1, pp. 18–26, 2014.

[8] M. Mon-Williams, J. P. Warm, and S. Rushton, "Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display," *Ophthalmic and Physiological Optics*, vol. 13, no. 4, pp. 387–391, 1993.

[9] I.-B. Pavaloiu, S.-A. Sandu, S.-D. Grigorescu, and G. Dragoi, "Inexpensive dentistry training using virtual reality tools," in *Proceedings of the 10th International Technology, Education and Development Conference, Valencia, Spain*, 2016, pp. 7–9.

[10] I. De Boer, P. Wesselink, and J. Vervoorn, "Student performance and appreciation using 3d vs. 2d vision in a virtual learning environment," *European Journal of Dental Education*, vol. 20, no. 3, pp. 142–147, 2016.

[11] L. M. Al-Saud, F. Mushtaq, I. Mirghani, A. Balkhoyor, A. Keeling, M. Manogue, and M. A. Mon-Williams, "Drilling into the functional significance of stereopsis: the impact of stereoscopic information on surgical performance," *Ophthalmic and Physiological Optics*, vol. 37, no. 4, pp. 498–506, 2017.

[12] E. Collaço, E. Kira, L. H. Sallaberry, A. C. Queiroz, M. A. Machado, O. Crivello Jr, and R. Tori, "Immersion and haptic feedback impacts on dental anesthesia technical skills virtual reality training," *Journal of Dental Education*, vol. 85, no. 4, pp. 589–598, 2021.

[13] C. Sutter, S. Sülzenbrück, M. Rieger, and J. Müsseler, "Limitations of distal effect anticipation when using tools," *New Ideas in Psychology*, vol. 31, no. 3, pp. 247–257, 2013.

[14] R. M. Wilkie, R. L. Johnson, P. R. Culmer, R. Allen, and M. Mon-Williams, "Looking at the task in hand impairs motor learning," *Journal of neurophysiology*, vol. 108, no. 11, pp. 3043–3048, 2012.

[15] Epic Games Inc. (2021) Digital Humans – metahuman creator. [Online]. Available: https://www.unrealengine.com/en-US/digital-humans

[16] 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)*, 2020, pp. 920–926.

[17] D. Kane, P. Guan, and M. S. Banks, "The limits of human stereopsis in space and time," *Journal of Neuroscience*, vol. 34, no. 4, pp. 1397–1408, 2014. [Online]. Available: https://www.jneurosci.org/content/34/4/1397