2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)

2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) | 979-8-3503-2838-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISMAR59233.2023.00134

Is this the vReal Life? Manipulating Visual Fidelity of Immersive Environments for Medical Task Simulation

Danny Schott
University of Magdeburg

Florian Heinrich
University of Wuerzburg

Lara Stallmeister
University of Magdeburg

Julia Moritz
USE-Ing. GmbH

Bennet Hensen
Hannover Medical School

Christian Hansen*
Universitity of Magdeburg

Figure 1: Virtual environment for evaluating an interventional procedure at different *visual fidelities* using controller and hand input. Left: Placement of a coil with tracked 3D printed replicas and hand interaction at the lowest level. Right: Performing the placement task of virtual objects using controller at medium level. Center: Highest level of *visual fidelity* as seen by the user.

ABSTRACT

Recent developments and research advances contribute to an ever-increasing trend towards quality levels close to what we experience in reality. In this work, we investigate how different degrees of these quality characteristics affect user performance, qualia of user experience (UX), and sense of presence in an example medical task. To this end, a two-way within-subjects design user study was conducted, in which three different levels of *visual fidelity* were compared. In addition, two different interaction modalities were considered: (1) the use of conventional VR controllers and (2) natural hand interaction using 3D-printed, spatially-registered replicas of medical devices, to interact with their virtual representations. Consistent results indicate that higher degrees of *visual fidelity* evoke a higher sense of presence and UX. However, user performance was less affected. Moreover, no differences were detected between both interaction modalities for the examined task. Future work should investigate the discovered interaction effects between quality levels and interaction modalities in more detail and examine whether these results can be reproduced in tasks that require more precision. This work provides insights into the implications to consider when studying interactions in VR and paves the way for investigations into early phases of medical product development and workflow analysis.

Index Terms: Human-centered computing—Virtual reality; Human-centered computing—User studies;

1 INTRODUCTION

Usability and user experience (UX) research in medical use cases, such as magnetic resonance imaging (MRI or, for short, MR) guided interventions, has the potential to improve patient outcomes and enhance the effectiveness of medical procedures. However, conducting such research faces a number of problems. Access during real procedures is limited as a result of ethical concerns, evaluation of technical devices is often difficult because of electromagnetic radiation, and the devices themselves are often fully occupied. Despite these challenges, it is crucial to consider the ergonomics and usability of interventions in the early development phases of prototypes to ensure optimal performance. To overcome these issues, a combination of physical mock-up and virtual reality (VR) hardware can be utilized to simulate MR interventions in virtual radiology suites [30, 38]. This approach allows for testing of early prototypes, drawing valuable conclusions that can be applied in real-world scenarios without requiring access to an MR device [34, 54]. Recent developments in VR technology enable compelling simulations and practice scenarios. The term realism in virtual environments (VE) describes the degree to which a simulation resemble our perception of reality we perceive every day in the real world [12]. Various studies have shown that these aspects of VR can have effects on the UX [7], perception [51] and on the sense of presence [12]. Realism is attained by incorporating multisensory stimuli, including visual, auditory, tactile, or agency indications [12]. Realistic lighting, along with dynamic shadows and reflections of moving objects, enhances these effects and has a positive impact on *visual fidelity* and, consequently, presence [19, 26, 53]. VR experiences are also influenced by the realistic nature and availability of input modalities. For instance, controllers are commonly used for interaction but lack flexibility and tactile feedback, making their use less realistic. Another potential

---

*e-mail: hansen@isg.cs.uni-magdeburg.de

DOI 10.1109/ISMAR59233.2023.00134

way to improve the realism of simulations is to represent interactive virtual objects with tangible physical mock-ups, e.g., by 3D-printing them and using trackers to determine their position and rotation in VR space [45]. Hand tracking can also lead to a more positive perception of the experience [50]. The combination of mixed reality visualization and tangible physical objects was shown to be more advantageous with respect to realism and enjoyment compared to either method alone [23]. In practice, the question arises as to which level of realism is needed to achieve desired effects on presence, UX or other related aspects. The amount of graphical detail may divert the developers' working time from optimizing interactions and the actual content of the simulation, in addition to increasing hardware requirements. For instance, Harman et al. [10] showed that minimal environments may suffice for memory recall tasks and Schmied-Kowarziki and Paelke [37] found that the level of environmental detail has only a nominal effect on acrophobia patients in virtual exposure therapy.

This work seeks to identify an appropriate level of *visual fidelity* and input modalities in medical task simulation, aiming to provide design guidelines within the context of HCI investigations during prototypical development processes. The example task of placing a radiological coil as part of a needle-based MR-guided procedure is simulated in terms of workflow and interaction. An experiment was conducted by manipulating *visual fidelity* and interaction modalities. This was aimed at creating a foundation and collecting design recommendations for future research activities in medical task simulation. Three different levels of *visual fidelity* were defined and an interaction approach using tangible objects was compared to traditional controller input. Effects of these manipulations on presence and UX as central quality measures for VR simulations were investigated. Also of interest was the influence of *visual fidelity* on user performance.

2 RELATED WORK

The use of VR in a medical context enables the acquisition of anatomical knowledge and the learning and training of medical skills while minimizing risks to the patient [4, 29, 32, 33, 35]. In the field of medical interventions, Van Nguyen et al. present a training system for performing biopsies in VR [49]. Snarby et al. presented a system that enables training of medical procedures by incorporating real-time image data [43]. The use of immersive VR in training environments can also improve the preparation of assisting personnel during surgeries [8]. The results of a study by Nakarada-Kordic et al. [27] suggest that a VE has the potential to improve patients' experience and prepare them for examinations compared with a simulated MRI scan. They also emphasize the benefits of VR in MRI examinations, as it can be a cost-effective and low-risk tool for knowledge transfer to both patients and physicians. However, this study does not use a mock-up or tangible objects in their VE. The sense of presence could be improved by including tactile feedback from furniture and interactive elements. In addition, the study does not compare the effects of different level of *visual fidelity* and does not provide a fully immersive simulation. Gonçalves et al. [7] most recently presented a review article on studies investigating the impact of realism in VR. They determined, that in general, realism positively influences UX. They also categorized other research results in several independent variables that have been manipulated to achieve different realism levels. The resulting categories were avatar visual, environment visual, audio, haptic, and olfactory content variables, as well as audio, haptic, interaction, camera, lights, and physics system variables. As our work primarily focuses on the visual aspects of VR environments and interaction modalities, the articles reviewed within it, primarily aligning with these categories, are considered relevant to our research objectives. The effects of geometric realism, i.e., polygon count and texture resolution, were investigated by Hvass et al. [11] in the field of VR video games. Higher degrees of realism evoked higher subjective presence ratings as well as stronger physiological fear responses. Newman et al. [28] conducted two experiments to examine the effects of environmental representation on perception. The first study found that VR experiences elicited more positive responses than watching videos, but were less immersive than real-life observations. In their second experiment, they examined the impact of VE realism on stress recovery and found that higher realism enhanced the process and increased the sense of presence. A comparative study by Mizuho et al. [24] investigated the effects of switching between virtual and real environments on memory. It was found that the visual quality of VEs had no impact on context-dependent forgetting and source-monitoring errors. Additionally, the study found that a high-fidelity VE significantly enhanced the feeling of presence. In the field of prototype usability testing, Zhou and Rau [55] compared a tangible physical mock-up of the prototype to a condition without haptic feedback. Visual output was created using either immersive VR or a monitor. The physical mock-up could improve performance, and VR generally evoked more positive subjective feedback. However, in the VR condition, using the tangible object did not improve *Involvement*. McMahan et al. [21] compared a natural interaction technique using tracked hand-held devices to traditional mouse and keyboard input in a VR game and additionally varied between a stereoscopic 360° CAVE display and a monoscopic single-wall display of that CAVE. They showed that the combination of low visual and low interaction fidelity performed comparably to the combination of high visual and high interaction fidelity, and that both outperformed the other two factor level combinations.

3 MATERIALS

Two major areas in existing scientific works are consideration of the lighting system [6, 36, 42, 53, 56] and object design [48]. These works have identified individual factors of the visual design of a VR application that have an impact on the user in the virtual world. However, few of these works relate to the medical context. In reviewing the existing literature, we found that a comprehensive investigation of all previously studied design components is lacking. Therefore, we suspect that a combination of these factors could either mitigate or amplify their individual effects. We developed a VE with three different fidelity characteristics (See Fig. 1). In them we combine abstract and realistic visual components, the selection of which is based on findings from previous scientific work in this field. In doing so, we increase the visual quality analogously to the development effort of such VE´s and have defined three gradations: *Essential Detail Level (EDL)* – This level prioritizes minimal development effort, focusing solely on task-relevant objects in the environment while representing the rest schematically with 3D primitives; *Standard Detail Level (SDL)* – At this level, we follow the typical effort involved in creating 3D environments, including basic shapes, materials, and lighting; *Advanced Detail Level (ADL)* – The primary emphasis at this level is on achieving optimal graphic implementation, with a strong focus on attaining a high level of visual accuracy to closely resemble realism. The detailed technical realization can be seen in Table 1. For our VR environment, we extracted the coil placement as an exemplary task from the interventional workflow because this action has a large range of motion and requires direct interactions with tangible objects at different locations in the room. The task reproduced in reality and its virtual translation can be seen in Fig. 2.

To provide users with the feeling of being in a real operating room (OR), we have added an audio recording of the pumping sound (compression of helium) made by an MRI scanner during surgery[1]. This sound is very loud, which is why headphones are worn and verbal communication is difficult. We consider this background noise to be an ambient sound that enhances the quality of the application. Therefore, we have added 2D audio in *SDL* and 3D (Spatial) audio

---

[1] (CC BY 3.0): freesound.org/people/solidphase/sounds/442831/

in *ADL*. To maintain minimal development effort in *EDL*, the audio was completely omitted. Under the same consideration regarding development effort, we also implemented different qualities of representation for a male and female hand. The difference was in *ADL*, where in addition to the photo-realistic hand model, a forearm was added. The gender feature was omitted in *EDL*.

As a realistic input modality, non-contact hand tracking is combined with real tracked objects to provide haptic and visual hand feedback while offering many degrees of freedom. A similar form of interaction was implemented by Schrom-Feiertag et al. [40]. A study by Bolder et al. [2] shows that implementing a form of interaction with tracked hands and real objects for interaction can achieve similar usability as in the real world. The controller represents a second input modality that allows the user to grasp and move virtual objects by pressing buttons. The implemented input modalities (controllers / hands-only) allow for the interaction with virtual replicas of tangible objects (coil, adapter, door handle) to perform the medical task at different *visual fidelities*.

4 EVALUATION

We conducted a study to evaluate the levels of *visual fidelity* and the interaction modalities of an exemplary medical task in a virtual OR in terms of their impact on presence, UX, and user performance. The following describes its methodology. The experiment followed a two-way within-subjects design that combined the different *visual fidelities (EDL, SDL, ADL)*, with the two *interaction modalities controller* and *tangible objects*. Because we wanted to evaluate effects of different fidelity alterations on a holistic task simulation, we opted for varying multiple parameters at the same time. That way, we tried to obtain more general answers first.

Figure 2: Realization of subtasks in reality including custom 3D printed tangible objects (left) vs. the analog realization in VR with hand interaction in *ADL* (right).

4.1 Hypotheses

This study aimed to investigate several hypotheses. We generally presume that interacting with tangible objects rather than controllers is more natural, and, thus, more realistic. First of all, we had to ensure that our manipulations are perceived as intended by the subjects. Therefore, the first two hypotheses are:

H1.1: Higher *visual fidelity* evoke higher subjective realism.
H1.2: Interacting with tangible objects is perceived as more realistic than controller-based interaction.

Based on related work [11, 28], we speculate that *visual fidelity* also has a positive effect on presence in our case. The following hypotheses are made to verify this assumption:

H2.1: Presence is positively affected by increasing *visual fidelity*.
H2.2: Tangible objects elicit a greater sense of presence than controller interaction.

In their literature review, Gonçalves et al. [7] showed that higher degrees of realism most often also had positive effects on UX. However, since not all reported studies could show this effect, we want to investigate this aspect as well:

H3.1: Higher *visual fidelity* are associated with a better UX.
H3.2: Tangible objects improve UX compared to controller input.

Stevens et al. [44] found a moderate correlation between presence and user performance. We were curious if such a relationship can also be shown in our experiment. The final hypotheses are:

H4.1: Better user performance can be observed for higher realism.
H4.2: Participants perform better using the tangible objects than using the controller.

4.2 Sample design

The study did not require professional expertise, as we believe this was not necessary to perform the experimental tasks. However, to ensure that subjects had some basic medical background knowledge and were familiar with medical environments, only participants studying human medicine in their second semester and higher were invited. We recruited 19 medical students aged 21-31 years (median (Mdn) = 24) from our university. Fourteen of them reported female gender (male = 5). Each individual was compensated 30 euros. A rating scale from 1 (no experience) to 5 (very experienced) was available for assessing technological experience, which resulted in the following distribution: technology affinity (Mdn = 3), gaming experience (Mdn = 2), VR experience (Mdn = 2). More than half of the participants stated that they had a medium technological affinity. One participant rated a high technological affinity. The majority of participants reported little to no (>60 %) experience with gaming and VR, with no one rating their experience at the highest level (5). All participants in our evaluation reported being in their 5th semester or higher of medical school.

4.3 Task

A radiological coil placement task was considered for this study. Participants were placed in an interventional MRI suite and had to move to three randomized positions in the room (subtask 0). Once there, a cabinet needed to be located and opened (Fig. 2 1). Participants found a flexible loop coil inside of it and needed to take it (Fig. 2 2). The coil then had to be placed on top of an already prepared virtual patient (Fig. 2 3). A mark on the patient's skin showed where the coil needed to be positioned. Participants were instructed to align the center of the circle-shaped coil with this marking. Finally, the coil needed to be connected to a plug positioned at the MRI couch to end the task (Fig. 2 4).

Table 1: Breakdown of the different visual fidelities and their technical implementation

	Advanced (ADL)	**Standard (SDL)**	**Essential (EDL)**
No. of polygons	>100.000 faces	>10000 faces	<10000 faces
Object features	Realistic shape replication, High-poly models, Natural irregularities	Essential geometric details (bevels / fillets), Medium-poly models	Basic geometric shapes, Low-poly models, No details (e.g. no bevels / fillets)
Material	HDR-materials, High resolution texture maps, Realistic properties & colors	Monochrome, Reflections, No texture maps	Monochrome, (no reflections & texture maps)
Light & Shadow	HQ shadows, Baked lightmap, Ambient occlusion	No shadows, Default light setup (6 Sources)	No shadows, Two light sources
Room architecture	High-detail OR-facilities (e.g. sockets, baseboards), Natural irregularities	Essential OR-facilities (e.g. doors, windows, ...)	Waiver of facilities (no doors & windows ...)
Interior	Full OR-furniture & Accessories	Basic OR-furniture, (no small-scale equipment)	Mandatory elements for task (no additional furniture)

4.4 Variables

Two independent variables were investigated in this experiment. Both were given by the factorial design of the study. The first manipulated independent variable was the *visual fidelity* of the VE. The second one was the respective *interaction modality*. Seven dependent variables were observed in the experiment. First, the *time* to complete the task was taken. Measurements started when the participants opened the closet door and stopped when the coil was plugged in. In addition, the *placement deviation* of the final coil position was assessed. This was the distance between the center of the circular part of the coil and a target marked on the virtual patient. The igroup presence questionnaire (*IPQ*) [41] was used to estimate presence-related qualia. The IPQ consists of the four sub-scales *General Presence*, *Spatial Presence*, *Involvement* and *Experienced Realism*. All scales were considered as dependent variables for this study. *Experienced Realism* primarily served the purpose of a manipulation check, to assess whether the intended manipulations of realism were successful. Additionally, the short version of the User Experience Questionnaire (*UEQ*) [39] was used to get an estimate on the users' subjectively perceived UX quality. This short version consists of 10 items that cumulatively form an overall *UX* score. Finally, a few control measures were collected. Participants were asked to rank their experiences with video games and with VR, as well as their technical affinity, on a scale with five gradations. Furthermore, the time participants spent in VR before starting the actual coil placement task was measured. This *Adjustment Period* was the time required to walk to all predefined floor markers.

4.5 Apparatus

The study took place in a usability lab at our university, where a mock-up with dimensions similar to a real MRI device was present [15]. We 3D printed a loop coil, a closet handle and a coil connector. The coil was placed inside a closet in one corner of the room and the printed handle was glued to one door of that closet (see Fig. 2). A short section of a plastic tube was attached to the coil connector to mimic the cable that would take its place in real MRI suites. The closet door, the loop coil, and the connector were all tracked using *HTC Vive* trackers mounted on them. The calibration of the VR handhelds was performed manually, relying on visual feedback and considering the displayed virtual and physical objects. A human torso dummy made of Styrofoam represented an average patient in size and position and was placed on the MRI bench. Velcro on this dummy was used to attach the 3D printed loop coil replica, which also had Velcro dots on its bottom. The coil connector was placed on top of the dummy to be near the user when needed. To ensure the greatest possible freedom of movement, a ceiling-mounted cable management system was used for the VR system. The laboratory had a size of about 50 m^2. The virtual tracking space, in which the participant could move, had a size of approx. $5x5\ m$. The VE represents a faithful replica in terms of room size, as well as placement of furniture and objects in the room.

The hardware used included a Valve Index HMD with the Stereo IR 170 Evaluation Kit (Ultraleap Ltd.) hand tracking sensor attached to the front. In addition to using four Valve Index Basestations (2.0) and controllers from the same manufacturer as additional input modalities, tangible objects were equipped with Vive Trackers 3.0 for position and rotation tracking. The PC used had the following specifications: Intel Core i7-8700K 3.70GHz 6-core CPU, NVIDIA GeForce RTX 3090 Ti, 32 GB RAM, 512 GB PCIE 3.0 SSD. Our tangible objects were 3D printed from Ultimaker PLA and TPU 95A (flexible materials). In regard to software, we used Unity 2020.3.31 in C# with Microsoft Windows 10 Pro (build 19044). For the implementation of hand tracking, we used Ultraleap´s Hand Tracking v5.2.0 and Leap Motion Core Plugin v5.5.0 for Unity. Rendering of *EDL* and *SDL* was performed in Unity's built-in render pipeline. In *ADL*, the High Definition Render Pipeline (HDRP) version 10.8.1 was used with materials from the Sample Scene (Measured Material Library for HDRP[2]). We used realistic female and male hand models (later assigned to subjects according to gender) from the Unity Assetstore [3]). In *SDL* we made changes to the material, whereas in *EDL*, standard low poly models were used.

4.6 Procedure

After welcoming the study participants, informed consent was obtained and demographic data were collected. Subsequently, the subject received a brief introduction to the medical field of application and an overview of the study procedure. As training, each participant first started performing the task in reality (without VR equipment), which was demonstrated once by the study investigator. For this purpose, the interaction objects and the spatial laboratory conditions (mock-up as a model for the interventional MRI) were addressed. The participant was reminded of the correct order in the interaction task: specifically, the order of coil placement and plug insertion was emphasized. The coil should be placed centrally over the target marker on the patient. Care should be taken to ensure

---

[2] https://github.com/Unity-Technologies/MeasuredMaterialLibraryHDRP

[3] assetstore.unity.com/packages/3d/characters/human- oids/leap-motion-realistic-female-hands-211090, 2022

Table 2: Summary of the ANOVAs' results on *spatial presence*, *involvement*, and *Experienced Realism*; as well as robust ANOVAs' results on *General Presence*, *UX*, *time*, and *placement deviation* ($\alpha < .05$). Test statistic F and effect size η^2 are reported for ANOVAs and test statistic Q and effect size δ_t are reported for robust ANOVAs.

Variable / Effect type	Factor	DFn	DFd	F/Q	p	Sig.	η^2/δ_t	Effect
Involvement								
Main effect(s)	*Visual fidelity*	2	34	16.541	<0.001	*	0.158	Large
Experienced Realism								
Main effect(s)	*Visual fidelity*	2	34	35.048	<0.001	*	0.372	Large
Interaction effect	*Fidelity * Interaction*	2	34	4.747	0.015	*	0.053	Small
General Presence								
Main effect(s)	*Visual fidelity*	2		10.581	<0.001	*	0.648	Medium
Interaction effect	*Fidelity * Interaction*	2		5.736	0.003	*	-	-
UX								
Main effect(s)	*Visual fidelity*	2		17.924	<0.001	*	0.595	Medium
Placement deviation								
Main effect(s)	*Visual fidelity*	2		9.845	<0.001	*	0.294	Small
	Interaction modality	1		4.984	0.025	*	0.466	Small
Interaction effect	*Fidelity * Interaction*	2		38.532	<0.001	*	-	-

that the Velcro side of the coil was faced down. The participant should perform the task at a reasonable pace and was not motivated to complete the task as quickly as possible. After training, the experimental task described in Sect. 4.3 was performed three times for each combination of *visual fidelity* and *interaction modality* factor level combination in VR. The order of the six resulting experimental conditions was partially randomized. *Visual fidelity* was randomly arranged for each participant. Both interaction modalities were tested one after the other for each respective *visual fidelity*. The order of the modalities was alternated between two participants. In *SDL* and *ADL*, corresponding male or female hand models were set, taking into account the gender indicated by the participants. During this period, participants needed to walk to four specific markings on the floor in a given order. Three different but comparable routes were predefined. This phase was included to allow the subjects to become accustomed to, and aware of, the environment before the actual task began. It was also performed during the training phase in reality, where tape was used to mark the positions on the floor.

4.7 Statistical analysis

The data for each dependent measure was first checked for homogeneity with Levene's tests. Next, normality assumptions were verified. To this end, the data of each variable was fitted to a linear model. Shapiro-Wilk tests were then conducted using the respective linear models' residuals to check for normality. In case homogeneity and normality assumptions were met, two-way repeated measures analyses of variance (ANOVAs) were conducted to analyse the data. For variables violating one of the assumptions, robust two-way ANOVAs for within-subject designs based on trimmed means were calculated to evaluate main and interaction effects (also see [52]). The δ_t estimate proposed by Algina et al. [1] was interpreted as effect size for main effects assessed by robust ANOVAs. Afterwards, post-hoc tests on statistically significant *visual fidelity* main effects were conducted. Pairwise paired t-tests with Bonferroni adjustments were applied to variables having met the normality assumption and robust Yuen's trimmed means tests with p-value adjustments using Hochberg's method were performed otherwise [52].

5 RESULTS

This section presents the complete set of experimental results. In terms of statistical outcomes, we began by calculating the mean values to aggregate placement deviation and time data under identical experimental conditions. Then, final *IPQ* sub-scale and *UEQ* scores were calculated according to the questionnaires instructions. Table 3 summarises all resulting descriptive results. No conducted Levene's test showed significant results. Thus, it was assumed that the homogeneity assumption held true for all dependent variables. However, significant Shapiro-Wilk test results on *General Presence*, *UX*, *time*, and *placement deviation* suggested that this data would not be normally distributed. Therefore, robust ANOVAs were conducted on these variables. *Spatial presence*, *Involvement*, and *Experienced Realism* were evaluated using conventional two-way repeated measures ANOVAs. Results of all these analyses are summarized in Table 2. Statistically significant *visual fidelity* main effects on *Experienced Realism*, *Involvement*, *General Presence*, and *UX* were revealed. These effects are visualized in Fig. 3. Pairwise comparison results are depicted in these plots. *Placement deviation* showed significant main effects for both factors. However, these are challenged by a significant interaction effect on that variable. *Placement deviation* results are visualized in Fig. 5. Significant interaction effects were also shown for the *Experienced Realism* and *UX* variables. These effects are visualized in Fig. 4.

5.1 Control measures

The participants' answers regarding video game experience, VR experience, and technological affinity were not evenly distributed. For example, 10 of 19 participants reported having a medium level of technical affinity, 12 of 19 participants had very little to little experience with video games and 11 of 19 participants said they had very little to little experience with VR. Because of this uneven distribution of group sizes, extensive statistical analyses of the effect of these measures did not seem viable. However, as exemplary analyses, Pearson's correlation tests between technical affinity and *Experienced Realism* were conducted for each combination of factor levels. No test returned significant results. Therefore, these control measures were not considered any further. Regarding the *Adjustment Period*, a Shapiro-Wilk test revealed a violation of normality. Therefore, this measure was analyzed with a robust two-way repeated measures ANOVA. This test showed a statistically significant *visual fidelity* main effect ($Q = 16.94, p < 0.001$). Follow-up pairwise comparisons using robust Yuen's trimmed means tests showed that participants spent significantly more time in the adjustment phase in *ADL* ($M = 7.75s, SD = 1.01s$) compared to both, *SDL* ($p = 0.001$; $M = 6.93s, SD = 0.85s$) and *EDL* ($p = 0.001$; $M = 6.89s, SD = 0.80s$).

5.2 Interpretation of results

The following attempts to interpret the identified effects and to find reasons for their occurrence. The *Experienced Realism* sub-scale of the *IPQ* questionnaire was considered as a manipulation check to ensure our conditions had the intended effects. Statistically significant differences between factor levels of the *visual fidelity* variable confirmed a clear ranking between conditions (see Fig. 3a). Therefore, we accept H1.1. However, no significant differences could

Table 3: Summary of descriptive results for all dependent variables ($n = 19$). All entries are in the format: mean value [standard deviation]. SP - *spatial presence*, INV - *involvement*, REAL - *experienced realism*, G - *general presence*, UX - *user experience*.

Variable	**SP**	**INV**	**REAL**	**G**	**UX**	**Time [s]**	**Deviation [mm]**
Accumulated	4.63 [0.60]	4.43 [1.22]	4.02 [1.22]	5.27 [1.25]	5.45 [1.17]	20.71 [6.79]	10.53 [6.32]
Advanced (*ADL*)	4.65 [0.58]	4.93 [1.12]	4.79 [0.87]	5.97 [0.88]	5.97 [0.95]	20.92 [6.87]	10.56 [5.57]
Standard (*SDL*)	4.62 [0.55]	4.58 [1.03]	4.22 [1.01]	5.25 [1.25]	5.56 [1.00]	20.55 [6.24]	9.00 [5.44]
Essential (*EDL*)	4.63 [0.67]	3.78 [1.24]	3.04 [1.07]	4.58 [1.20]	4.83 [1.25]	20.67 [7.40]	12.03 [7.53]
Tangible	4.63 [0.62]	4.33 [1.28]	4.06 [1.34]	5.31 [1.24]	5.43 [1.21]	20.74 [6.88]	11.37 [6.47]
Advanced (*ADL*)	4.64 [0.58]	4.86 [1.26]	4.86 [0.88]	5.94 [0.73]	5.96 [1.03]	20.58 [6.71]	9.60 [3.89]
Standard (*SDL*)	4.63 [0.60]	4.61 [1.03]	4.51 [1.04]	5.67 [0.97]	5.61 [1.02]	20.91 [6.80]	7.96 [3.47]
Essential (*EDL*)	4.61 [0.71]	3.51 [1.17]	2.79 [1.08]	4.33 [1.33]	4.72 [1.26]	20.72 [7.48]	16.54 [7.71]
Controller	4.63 [0.58]	4.53 [1.16]	3.98 [1.10]	5.22 [1.27]	5.47 [1.13]	20.69 [6.77]	9.70 [6.11]
Advanced (*ADL*)	4.66 [0.61]	5.00 [1.00]	4.72 [0.88]	6.00 [1.03]	5.97 [0.89]	21.27 [7.20]	11.51 [6.84]
Standard (*SDL*)	4.60 [0.51]	4.54 [1.06]	3.92 [0.90]	4.83 [1.38]	5.51 [1.01]	20.18 [5.80]	10.05 [6.82]
Essential (*EDL*)	4.64 [0.64]	4.06 [1.28]	3.29 [1.04]	4.83 [1.04]	4.93 [1.27]	20.61 [7.54]	7.52 [3.78]

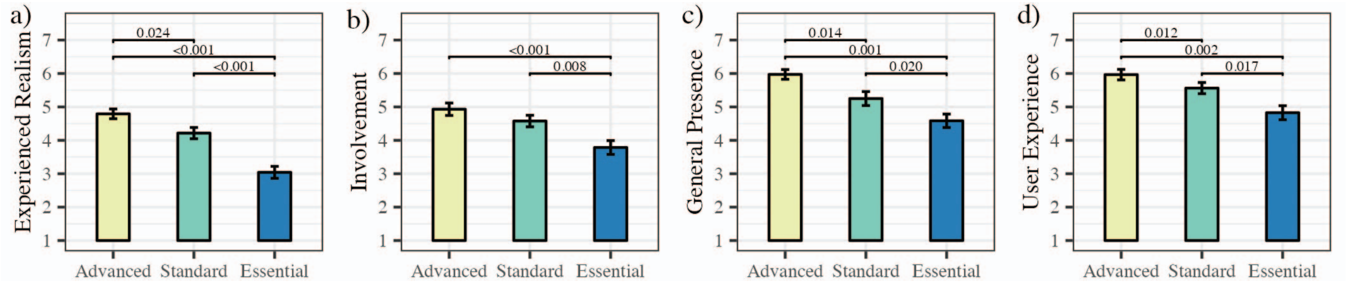

Figure 3: Significant main effects of the *visual fidelity* factor on: a) *Experienced Realism*, b) *Involvement*, c) *General Presence*, and d) *UX*. Bars represent mean values and error bars represent standard errors. Significant post-hoc test results are highlighted with brackets.

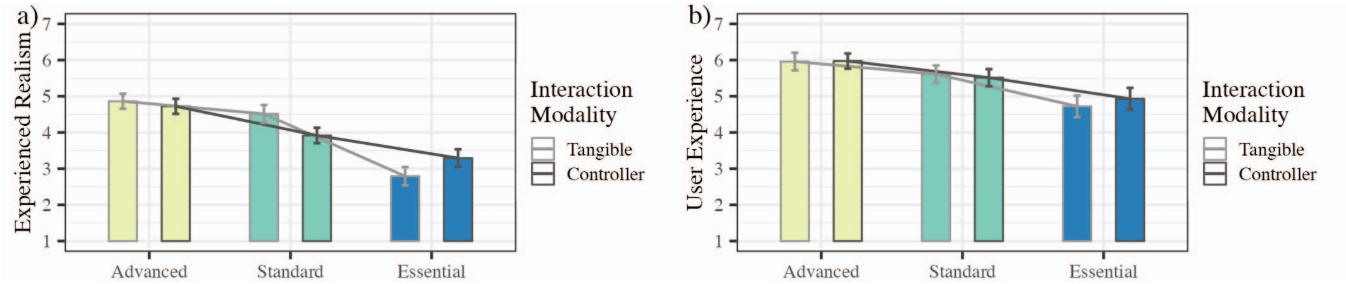

Figure 4: Significant interaction effects: a) *Experienced Realism*, and b) *UX*. Bars represent mean values and error bars represent standard errors. Lines connect *visual fidelity* mean values with respect to each *interaction modality* factor to visualize the interaction effects.

be detected regarding the *interaction modality* factor. Observations during the study and inspection of the raw data revealed that whether tangible devices or controllers were perceived more realistically seemed rather user dependent. Thus, we cannot accept H1.2.

In addition, a significant interaction effect was shown regarding the *Experienced Realism* variable (see Fig. 4a). The tangible interaction modality seemed more realistic than the controller modality in *SDL*. However, this assessment was reversed in *EDL*, which is why this significant interaction effect occurred. Because the higher *Experienced Realism* assessment in *EDL* (using the controller) was still below the smaller value of in *SDL* (also using the controller), we argue that the significant *visual fidelity* main effect is valid despite the interaction effect and does not influence our acceptance of H1.1.

A reason for the interaction may be found in the visualization of the interaction devices. The *EDL* showed very abstract hand models without textures, while *SDL* showed already quite realistic ones. At the same time, the controller's appearance did not change much across conditions. It may be plausible that participants focused more on their hands when operating tangible objects than when using the controller. Therefore, the more realistic-looking controller evoked a higher degree of *Experienced Realism* than the abstract hands in *EDL*. At the same time, the more realistic hands in *SDL* may have been perceived as more natural and, thus, more realistic, than the controller. Interestingly, a comparable effect could not be shown for the *ADL* condition. Here, both interaction modalities ranked very similarly. This may have been because the inherent *visual fidelity* of the VE caused participants to focus more on their surroundings and less on their hands. Thus, when filling in the *IPQ* questionnaire after VR exposure, they ranked both interaction modalities similarly for this *visual fidelity*. An indicator for this can be found in the *Adjustment Period* control measure. Participants required significantly more time to navigate all the ground markings in *ADL*. We argue that this was caused by the subjects spending more time observing their surroundings and, as a result, concentrating less on the actual task. Consistent results of the *visual fidelity* factor show a clear ranking and relationship between the degree of visual quality and VR-related qualia. Higher *visual fidelity* resulted in a higher sense of ”being there” (Fig. 3c), caused participants to feel more involved in the VE and to devote more attention to it (Fig. 3b), and to have a generally better UX (Fig. 3d). Therefore, we consider H2.1 and H3.1 as accepted.

Regarding the latter, half of the items of the *UEQ* are related to the hedonic quality of the system. Hence, it was less surprising that improving the visual quality and attractiveness of our VE was also reflected in this measure. However, we believe that confirming this hypothesis is an important research finding. The significant interaction effect on *UX* is similar to the one on *Experienced Realism*. Likewise, it can be explained analogously and was probably caused by the controller being rendered similarly in *EDL* and *SDL*, while the respective hand models showed visual differences. These observed differences are comparably low and the overall trend between *visual fidelity* conditions seems unaffected by this interaction.

The increased *Involvement* ratings constitute a more interesting finding. Participants were less aware of their surrounding real world in *ADL*. This indicates that the increasing match between the VE and the real world blurred the boundaries between both realities. However, we do not know if this was caused by visual quality improvements or by the increased amount of observable items in the VE at higher *visual fidelity*. The *General Presence* item of the *IPQ* questionnaire was answered significantly different between *visual fidelities*. It is said to be closely related to *Spatial Presence* [41]. However, this sub-scale did not seem to be affected by the *visual fidelity* factor in our study, which partly diminishes our acceptance of H2.1. The *General Presence* item asked participants if they had a sense of ”being there”. This question could have been interpreted ambiguously and other aspects, e.g., *Experienced Realism* and *Involvement*, could have also affected responses. *Spatial Presence*, as in the sense of being *physically* present in the VE, may be more susceptible to other factors (such as display-related properties) that were not manipulated in our study.

Except for *placement deviation*, which will be discussed separately, no statistically significant differences were found between both interaction modalities in this work. Since the descriptive data also does not reveal any meaningful insights, we argue that, for the selected task, the choice of interaction modality should be rather user preference-driven. We therefore reject the hypotheses H2.2 and H2.3. We expected that interacting with haptic tangible objects would evoke a more positive UX response. Its absence may have been due to implementation reasons. All tangible objects were tracked using only one tracker. In addition, some parts were flexible, which could not be translated to VR. Visuoproprioceptive mismatches between the displayed virtual objects and their real world counterparts caused by tracking or registration errors may have negatively affected participants’ perception of this modality.

Two user performance measures were considered. The time variable showed no significant differences. In contrast, two significant main effects were found on coil *placement deviation*. However, both appear to be strongly affected by a significant interaction effect on this variable. In *SDL* and *ADL*, participants were more accurate using the tangible objects. In contrast, subjects could place the coil with a similar accuracy using the controller in the *EDL* but performed worst using the tangible objects in this *visual fidelity*. Therefore, we are not able to accept H4.1 and H4.2. Considering a loop coil diameter of 18.5 *cm*, observed differences in deviations below 1 *cm* seem only minor. Nonetheless, the statistically significant nature of these differences is noteworthy. Perhaps participants were the most accurate overall using the controller in *EDL*, because this factor level combination introduced the least amount of distracting stimuli, thus providing an environment for more concentrated work. In the more realistic VEs, participants may have benefited from the more natural tangible interaction modality creating an overall workflow close to reality that facilitated more concentrated performance.

6 GENERAL DISCUSSION

We identified significant effects from the degree of *visual fidelity* on several subjective measures. This research finding is in line with the work of Newman et al. [28] and Mizuho et al. [24], who also found that a high-quality VE created a greater sense of presence. The interaction modality did not seem to affect presence-related qualia in our study, which is similar to the work of Zhou and Rau [55]. There, using a tangible object in VR did also not improve *Involvement* compared to a standard condition.

Our experiment’s sample size was relatively small. A sensitivity power analysis showed that, with a β of 0.8, effects up to a Cohen’s f effect size of 0.245 were correctly identified with reasonable probability. This correlates to medium and large sized effects. Hence, we cannot be certain to have potentially missed small-sized effects.

A higher proportion of female participants took part in our evaluation. Our selection process prioritized availability and willingness to participate, with no gender-specific criteria being influential. As we recruited knowledgeable medical students in general, we came up with a random sampling approach that could potentially impact the generalizability of our findings. As a result, future research should aim to extend our findings to broader populations to improve the external validity of the study and ensure inclusion and diversity.

The virtual hands of the user were rendered differently for each *visual fidelity*. This may have had uncontrolled effects on the virtual body ownership (VBO) of participants and may therefore have affected the results. In the literature, VBO is assumed to be evoked by a combination of coherent sensorimotor stimuli (bottom-up), e.g., moving one’s real hand results in a analog behavior of the virtual hand, and congruencies on the cognitive layer (top-down), e.g., visual resemblance of the virtual hand and one’s own hand [47]. All conditions in our study implemented the same hand tracking mechanisms and, thus, did not differ on the sensorimotor bottom-up layer. Maselli & Slater [20] concluded that coherence on this layer induces vivid VBO, even for mannequin-resembling virtual hands. Lugrin et al. [18] also showed that non-human avatars can elicit comparable, and even slightly higher, degrees of VBO and concluded that realistic appearances are not critical top-down factors. In contrast, an experiment by Latoschik et al. [17] revealed that photo-realistic avatars evoke a stronger acceptance of the virtual body as one’s own compared to a wooden mannequin. Therefore, it is uncertain whether the manipulation of the virtual hands’ appearance had an effect in our study. This should be investigated in the future.

We observed interaction effects on *visual fidelity*, *UX*, and *placement deviation*. *EDL* was rated higher and performed better when paired with controller interaction. *ADL* received higher ratings and exhibited lower placement deviations when tangible objects were utilized. The reversed factor level combinations may have then performed worse because the mismatch caused an uncanny valley-like effect [25]. McMahan et al. [22] investigated whether such an effect also applies to interaction in VR. They theorize that, as with robots, more natural interaction paradigms will feel and perform worse after a certain degree and will only improve once a high resemblance to real world interaction is achieved. Both the controller and the tangible interaction modalities performed very similarly in our study. The tangible condition was designed to feel more natural and to be, thus, more realistic. However, the final prototypical implementation may have just not been good enough. Slight tracking and registration errors may have caused a feeling of eeriness that is associated with the uncanny valley effect.

In general, *Experienced Realism* in *ADL* received lower ratings, as we would have expected ($M = 4.65$ on a scale from 1 to 7). This suggests that there is still some room for improvement. Future work could investigate if different VR headsets have an impact on this measure. For example, the Varjo XR-3 (Varjo Technologies Oy, Helsinki, Finland) was evaluated to provide a very high visual acuity [14]. Moreover, the visual realism of our VE could be improved

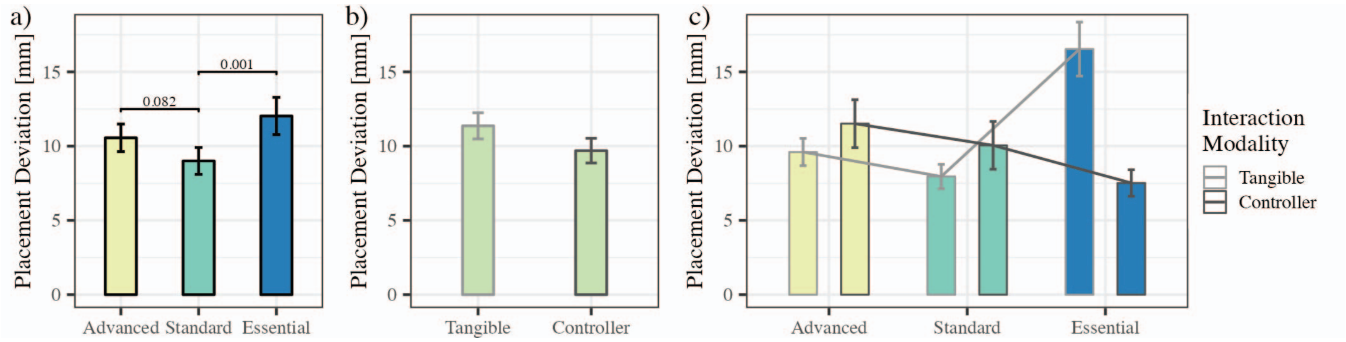

Figure 5: Significant main and interaction effects on *placement deviation*. a) *visual fidelity main effect*, b) *interaction modality main effect*, and c) interaction effect. Bars represent mean values and error bars represent standard errors. Significant post-hoc test results are highlighted with brackets. Lines connect *visual fidelity* mean values with respect to each *interaction modality* factor to visualize the interaction effect.

by including social companions. Previous research showed that the quality of experience can be positively affected by co-locating an increasing amount of such virtual agents [16]. Therefore, having radiology technologists, anesthesiologists or surgical nurses joining the VE could enhance the *Experienced Realism*.

A major limitation of our study is that we cannot draw meaningful conclusions on which specific *visual fidelity* aspect was responsible for the observed effects (see Table 1). For this first experiment, we decided instead to compare holistic impressions. Multiple follow-up studies would be required to find answers as to which individual components had stronger effects. We used appropriate 3D hand models in *SDL* and *ADL* based on the indicated gender of the participants. Personalized visual body representations with a fully represented arm elicit a high sense of virtual body ownership [13]. The absence of these features (forearm connection) in *EDL* and *SDL* may have influenced our results, as this effect is facilitated by dynamic actions [46]). Hand models exhibit varying sizes at least based on gender. However, the study did not account for variations in individual hand size and skin tone, potentially introducing factors that could influence interactions with objects. While our sample did not include participants from diverse backgrounds, we acknowledge the significance of representing a broader range of skin colors within virtual environments to prevent potential biases and ensure inclusivity. In real life, gloves are worn to maintain sterility, but we refrain from using them in this study as it would further influence the illusion [9].

Following Fink et al. [5], we used obstacles in the form of primitive 3D objects as placeholders where real objects are located in *EDL*, which constrained locomotion in virtual space. This resulted in the inability to define specific medical devices, but it allowed us to focus the participants on the execution of the task. In addition, it remains uncertain if different tasks would have resulted in different outcomes. No differences were identified between interaction modalities in this study. However, tasks requiring more precision could potentially benefit from the more natural tangible interaction paradigm. Future work could therefore repeat the experiment with a different task design. Additionally, our sample design included only medical students because no interventional expertise was required. However, clinicians working in ORs or radiology suites on a daily basis may experience our digital twin of the MRI room differently because of their prior knowledge. Since they are already accustomed to the environment, they could potentially focus less on their surroundings at higher *visual fidelity*. At the same time, they could notice errors or mismatches that medical students would not see, which in turn could cause distractions. Hence, conducting a similar study with subjects from the expert domain would also be a meaningful continuation of this project. We used a realistic MRI background noise as ambient noise in *SDL* and *ADL* levels to enhance the feeling of being in an OR. In order to maintain minimal development requirements, we decided not to integrate sound in *EDL*. However, this may have caused *EDL* to score less well, as Dinh et al. [3] emphasize that the use of sound has a greater impact on presence than *visual fidelity*. The effect of contrast between presence and absence of sound at different visual levels is also supported by Poeschl et al. [31]. Since the audio used is very monotone and no additional audio effects (such as when interacting with objects) were included, we suspect that the influence of the use of audio regarding presence in our study will be very small.

7 CONCLUSION

This work investigated the effects of three different gradations of *visual fidelity* on presence and UX for a medical task in VR. Tow interaction modalities were investigated: The first one was based on natural hand interaction with tangible, 3D-printed objects and the second paradigm used conventional VR controllers. A control measure confirmed the successful manipulation of *Experienced Realism* with the three investigated *visual fidelities*. Our results revealed a strong connection between *visual fidelity* and the dependent measures of *General Presence*, *Involvement*, and UX. *Spatial Presence* and user performance were less affected. Moreover, no differences were detected between both interaction modalities for the examined task. Inspecting the raw data revealed that it seemed rather user dependent whether tangible devices or controllers were perceived more realistically. Future work should examine whether these results can be reproduced in tasks that require more precision. In addition, clinical experts should be included in follow-up studies and individual components of our considered *visual fidelities* could be investigated separately. The identified advantages of high *visual fidelity* in VEs have practical implications for the development of future systems. However, the consideration of development effort in future work might become redundant, as the manipulation of the level of detail using technologies such as AI-based filters and 3D model conversion could be accomplished without additional effort. Investigations of interaction paradigms in VR seem less dependent on the visual quality of the VE and do not seem to require high standards. However, experiments focused on VR related qualia, such as presence, can benefit from extensive efforts towards realism.

ACKNOWLEDGMENTS

This project received funding from the Federal Ministry of Education and Research as part of the Magdeburg Research Campus *STIMULATE* program (grant no. 13GW0473A), and from the German Federal Ministry for Economic Affairs and Climate Action (grant no. 16KN093943).

REFERENCES

[1] J. Algina, H. J. Keselman, and R. D. Penfield. An alternative to cohen's standardized mean difference effect size: A robust parameter and confidence interval in the two independent groups case. *Psychological Methods*, 10(3):317–328, 2005. doi: 10.1037/1082-989x.10.3.317

[2] A. Bolder, S. M. Grünvogel, and E. Angelescu. Comparison of the usability of a car infotainment system in a mixed reality environment and in a real car. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10, 2018.

[3] H. Q. Dinh, N. Walker, L. F. Hodges, C. Song, and A. Kobayashi. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, pp. 222–228. IEEE, 1999.

[4] R. Doerner, W. Broll, P. Grimm, and B. Jung. *Virtual and augmented reality (VR/AR): Foundations and methods of extended realities (XR)*. Springer Nature, 2022.

[5] P. W. Fink, P. S. Foo, and W. H. Warren. Obstacle avoidance during walking in real and virtual environments. *ACM Transactions on Applied Perception (TAP)*, 4(1):2–es, 2007.

[6] T. Fukuda, H. Ban, K. Yagi, and J. Nishiie. Development of high-definition virtual reality for historical architectural and urban digital reconstruction: A case study of azuchi castle and old castle town in 1581. In *Computer-Aided Architectural Design Futures. The Next City-New Technologies and the Future of the Built Environment: 16th International Conference, CAAD Futures 2015, São Paulo, Brazil, July 8-10, 2015. Selected Papers 16*, pp. 75–89. Springer, 2015.

[7] G. Gonçalves, H. Coelho, P. Monteiro, M. Melo, and M. Bessa. Systematic review of comparative studies of the impact of realism in immersive virtual experiences. *ACM Computing Surveys*, 55(6):1–36, 2022. doi: 10.1145/3533377

[8] A. Gupta, J. Cecil, and M. Pirela-Cruz. Immersive virtual reality based training and assessment of an orthopedic surgical process. In *2020 IEEE International Systems Conference (SysCon)*, pp. 1–7. IEEE, 2020.

[9] A. Haans, W. A. IJsselsteijn, and Y. A. de Kort. The effect of similarities in skin texture and hand shape on perceived ownership of a fake limb. *Body Image*, 5(4):389–394, 2008.

[10] J. Harman, R. Brown, D. Johnson, and S. Turkay. The role of visual detail during situated memory recall within a virtual reality environment. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*. ACM, 2019. doi: 10.1145/3369457.3369469

[11] J. Hvass, O. Larsen, K. Vendelbo, N. Nilsson, R. Nordahl, and S. Serafin. Visual realism and presence in a virtual reality game. In *2017 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*. IEEE, 2017. doi: 10.1109/3dtv.2017.8280421

[12] S. Jung and R. W. Lindeman. Perspective: Does realism improve presence in vr? suggesting a model and metric for vr experience evaluation. *Frontiers in Virtual Reality*, 2:693327, 2021.

[13] S. Jung, C. Sandor, P. J. Wisniewski, and C. E. Hughes. Realme: The influence of body and hand representations on body ownership and presence. In *Proceedings of the 5th Symposium on Spatial User Interaction*, pp. 3–11, 2017.

[14] E. Kappler, R. F. Jacinto, and S. Arndt. Evaluation of visual acuity and perceptual field of view using the varjo XR-3 headset in a virtual environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 66(1):2193–2197, 2022. doi: 10.1177/1071181322661195

[15] R. Kwapik, J. Moritz, B. Hensen, B. Janny, E. Pannicke, D. Schott, et al. Virtual reality-based usability laboratory for interventional mr applications. In *Proc 5th Conf on Image-Guided Interventions*, pp. 51–52, 2021.

[16] M. E. Latoschik, F. Kern, J.-P. Stauffert, A. Bartl, M. Botsch, and J.-L. Lugrin. Not alone here?! scalability and user experience of embodied ambient crowds in distributed social virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):2134–2144, 2019. doi: 10.1109/tvcg.2019.2899250

[17] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*. ACM, 2017. doi: 10.1145/3139131.3139156

[18] J.-L. Lugrin, J. Latt, and M. E. Latoschik. Anthropomorphism and illusion of virtual body ownership, 2015. doi: 10.2312/EGVE.20151303

[19] K. Maruya and T. Ohtani. Shadows can change the shape appearances of real and virtual objects. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1076–1077. IEEE, 2019.

[20] A. Maselli and M. Slater. The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience*, 7, 2013. doi: 10.3389/fnhum.2013.00083

[21] R. P. McMahan, D. A. Bowman, D. J. Zielinski, and R. B. Brady. Evaluating display fidelity and interaction fidelity in a virtual reality game. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):626–633, 2012. doi: 10.1109/tvcg.2012.43

[22] R. P. McMahan, C. Lai, and S. K. Pal. Interaction fidelity: The uncanny valley of virtual reality interactions. In *Lecture Notes in Computer Science*, pp. 59–70. Springer International Publishing, 2016. doi: 10.1007/978-3-319-39907-2_6

[23] X. Min, W. Zhang, S. Sun, N. Zhao, S. Tang, and Y. Zhuang. VP-Model: High-fidelity product simulation in a virtual-physical environment. *IEEE Transactions on Visualization and Computer Graphics*, 25(11):3083–3093, 2019. doi: 10.1109/tvcg.2019.2932276

[24] T. Mizuho, T. Narumi, and H. Kuzuoka. Effects of the visual fidelity of virtual environments on presence, context-dependent forgetting, and source-monitoring error. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2607–2614, 2023.

[25] M. Mori, K. F. MacDorman, and N. Kageki. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2):98–100, 2012. doi: 10.1109/MRA.2012.2192811

[26] J. Mortensen, I. Yu, P. Khanna, F. Tecchia, B. Spanlang, G. Marino, and M. Slater. Real-time global illumination for vr applications. *IEEE Computer Graphics and Applications*, 28(6):56–64, 2008.

[27] I. Nakarada-Kordic, S. Reay, G. Bennett, J. Kruse, A.-M. Lydon, and J. Sim. Can virtual reality simulation prepare patients for an mri experience? *Radiography*, 26(3):205–213, 2020.

[28] M. Newman, B. Gatersleben, K. Wyles, and E. Ratcliffe. The use of virtual reality in environment experiences and the importance of realism. *Journal of Environmental Psychology*, 79:101733, 2022. doi: 10.1016/j.jenvp.2021.101733

[29] J. Olasky, G. Sankaranarayanan, N. E. Seymour, J. H. Magee, A. Enquobahrie, M. C. Lin, R. Aggarwal, L. M. Brunt, S. D. Schwaitzberg, C. G. Cao, et al. Identifying opportunities for virtual reality simulation in surgical education: a review of the proceedings from the innovation, design, and emerging alliances in surgery (ideas) conference: Vr surgery. *Surgical innovation*, 22(5):514–521, 2015.

[30] E. F. Picka, A. Vogel, M.-S. Roder, J. Birkle, D. Schrenk, J. Linnemann, J. Moritz, and S. Pfeffer. Virtual usability testing (virtuse)-development of a methodical approach for usability testing in vr. In *HCI International 2022 Posters: 24th International Conference on Human-Computer Interaction, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings, Part I*, pp. 109–116. Springer, 2022.

[31] S. Poeschl, K. Wall, and N. Doering. Integration of spatial sound in immersive virtual environments an experimental study on effects of spatial sound on presence. In *2013 IEEE Virtual Reality (VR)*, pp. 129–130. IEEE, 2013.

[32] B. Preim and P. Saalfeld. A survey of virtual human anatomy education systems. *Computers & Graphics*, 71:132–153, 2018.

[33] B. Preim, P. Saalfeld, and C. Hansen. Virtual and augmented reality for educational anatomy. In *Digital Anatomy: Applications of Virtual, Mixed and Augmented Reality*, pp. 299–324. Springer, 2021.

[34] A. V. Reinschluessel, J. Teuber, M. Herrlich, J. Bissel, M. van Eikeren, J. Ganser, et al. Virtual reality for user-centered design and evaluation of touch-free interaction techniques for navigating medical images in the operating room. In *Ext Abstr Hum Factors Computing Syst*, pp. 2001–2009, 2017.

[35] R. Riener, M. Harders, R. Riener, and M. Harders. *Introduction to virtual reality in medicine*. Springer, 2012.

[36] R. Schatz, G. Regal, S. Schwarz, S. Suettc, and M. Kempf. Assessing the qoe impact of 3d rendering style in the context of vr-based training. In *2018 Tenth international conference on quality of multimedia experience (QoMEX)*, pp. 1–6. IEEE, 2018.

[37] R. L. Schmied-Kowarzik and V. Paelke. Examining the importance of realism in virtual reality therapy environments for people with specific phobias. In M. Weier, M. Bues, and R. Wechner, eds., *GI VR / AR Workshop*. Gesellschaft für Informatik e.V., 2021. doi: 10.18420/vrar2021_16

[38] D. Schott, F. Heinrich, D. Labsch, B. Hensen, and C. Hansen. Towards multimodal interaction for needlebased procedures in a virtual radiology suite. *Current Directions in Biomedical Engineering*, 8(1):70–73, July 2022. doi: 10.1515/cdbme-2022-0018

[39] M. Schrepp, A. Hinderks, and J. Thomaschewski. Design and evaluation of a short version of the user experience questionnaire (UEQ-s). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(6):103, 2017. doi: 10.9781/ijimai.2017.09.001

[40] H. Schrom-Feiertag, G. Regal, J. Puthenkalam, and S. Suette. Immersive experience prototyping: Using mixed reality to integrate real devices in virtual simulated contexts to prototype experiences with mobile apps. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 75–81. IEEE, 2021.

[41] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–281, 2001. doi: 10.1162/105474601300343603

[42] M. Slater, P. Khanna, J. Mortensen, and I. Yu. Visual realism enhances realistic response in an immersive virtual environment. *IEEE computer graphics and applications*, 29(3):76–84, 2009.

[43] H. Snarby, T. Gåsbakk, E. Prasolova-Førland, A. Steinsbekk, and F. Lindseth. Procedural medical training in vr in a smart virtual university hospital. In *Smart Education and e-Learning 2018 5*, pp. 132–141. Springer, 2019.

[44] J. A. Stevens and J. P. Kincaid. The relationship between presence and performance in virtual simulation training. *Open Journal of Modelling and Simulation*, 03(02):41–48, 2015. doi: 10.4236/ojmsi.2015.32005

[45] C. Taylor and D. Cosker. Interacting with real objects in virtual worlds. *Real VR–Immersive Digital Reality: How to Import the Real World into Head-Mounted Immersive Displays*, pp. 337–353, 2020.

[46] G. Tieri, E. Tidoni, E. F. Pavone, and S. M. Aglioti. Mere observation of body discontinuity affects perceived ownership and vicarious agency over a virtual hand. *Experimental brain research*, 233:1247–1259, 2015.

[47] M. Tsakiris and P. Haggard. The rubber hand illusion revisited: Visuotactile integration and self-attribution. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1):80–91, 2005. doi: 10.1037/0096-1523.31.1.80

[48] M. van Gisbergen, M. Kovacs, F. Campos, M. van der Heeft, and V. Vugts. What we don't know. the effect of realism in virtual reality on experience and behaviour. *Augmented Reality and Virtual Reality: The Power of AR and VR for Business*, pp. 45–57, 2019.

[49] D. Van Nguyen, S. B. Lakhal, and A. Chellali. Preliminary evaluation of a virtual needle insertion training system. In *2015 IEEE Virtual Reality (VR)*, pp. 247–248. IEEE, 2015.

[50] J.-N. Voigt-Antons, T. Kojic, D. Ali, and S. Möller. Influence of hand tracking as a way of interaction in virtual reality on user experience. In *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–4. IEEE, 2020.

[51] I. A. Wijayanto, S. V. Babu, C. C. Pagano, and J. H. Chuang. Comparing the effects of visual realism on size perception in vr versus real world viewing through physical and verbal judgments. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2721–2731, 2023.

[52] R. R. Wilcox. Chapter 8 - comparing multiple dependent groups. In R. R. Wilcox, ed., *Introduction to Robust Estimation and Hypothesis Testing*, pp. 467–539. Academic Press, fifth edition ed., 2022. doi: 10.1016/B978-0-12-820098-8.00014-2

[53] I. Yu, J. Mortensen, P. Khanna, B. Spanlang, and M. Slater. Visual realism enhances realistic response in an immersive virtual environment-part 2. *IEEE Computer Graphics and Applications*, 32(6):36–45, 2012.

[54] A. Zaman, A. Roy, K. Fatema, N. J. Farin, T. Doring, and R. Malaka. Explore voice and foot-based interaction techniques to navigate 2d radiological images in the virtual reality operation theatre. In *22nd Int Conf Comput Inf Technol*, pp. 1–7. IEEE, 2019.

[55] X. Zhou and P.-L. P. Rau. Determining fidelity of mixed prototypes: Effect of media and physical interaction. *Applied Ergonomics*, 80:111–118, Oct. 2019. doi: 10.1016/j.apergo.2019.05.007

[56] P. Zimmons and A. Panter. The influence of rendering quality on presence and task performance in a virtual environment. In *IEEE Virtual Reality, 2003. Proceedings.*, pp. 293–294. IEEE, 2003.