The effects of VR in training simulators: Exploring perception and knowledge gain

Aline Menin, Rafael Torchelsen, Luciana Nedel

Computers & Graphics (2021)

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

The Effects of VR in Training Simulators: Exploring Perception and Knowledge Gain

Aline Menin[a,*], Rafael Torchelsen[b], Luciana Nedel[c]

[a] *Univ. Côte d'Azur, CNRS, Inria, Sophia Antipolis, France*
[b] *CDTec, Federal University of Pelotas, Pelotas, Brazil*
[c] *Institute of Informatics, Federal University of Rio Grande do Sul, Porto Alegre, Brazil*

ARTICLE INFO

Article history:
Received October 12, 2021

Keywords: Virtual Reality, Immersion, Training Simulators, User Perception, Knowledge Gain

ABSTRACT

Although immersive virtual environments have been used for years for training and learning purposes (e.g., flight and surgery simulators), the effects of using VR devices on simulation sessions are yet to be understood. In this work, we explore the effects of different VR devices on virtual environments developed for training, focusing on perception and knowledge gain aspects. We performed two user studies to investigate the influence of these devices on users' workload, motion sickness, and performance in the domain of work safety training. The first experiment includes 61 participants and seeks to understand whether and how VR displays providing different fields of view affects the users' ability to search for risks in an office-like virtual environment (i.e. focus on user perception). Subsequently, we conducted a second experiment involving 46 subjects, where we assess whether and how interaction techniques providing different degrees-of-freedom influence users' ability to learn procedural tasks (i.e. focus on knowledge gain). From our results, we learned that users' knowledge on the simulation's topic (i.e. work safety) and gaming experience play an important role in VR simulations, and that cybersickness symptoms such as disorientation are likely caused by unawareness of one's surroundings instead of VR content.

1. Introduction

The low-cost and danger-free aspects of virtual reality (VR) technology enable the design of immersive simulations with game aspects (i.e. gameplay, challenge, interaction, and objective) for training people in diverse application domains. There are numerous successful examples on the literature on how immersive VR can improve user performance and game effectiveness. Chitarro et al. [1] proposed an immersive simulator for educating passengers on aviation safety through experiencing a serious aircraft emergency situation using a Head-Mounted Display (HMD). In fire service, Backlund et al. [2] present a simulator to train firefighter's skills using a Cave Automatic Virtual Environment (CAVE) based system and allowing free interaction through motion tracking. In health-care, Pedraza-Hueso et al. [3] propose a simulator that allows the user to carry out physical and cognitive rehabilitation therapies using a natural user interface based on motion tracking. Cecil et al. [4] proposed a VR simulator to facilitate and supplement the training opportunities provided to orthopedic residents. Mobach [5] showed that VR simulations can incite users to actively engage in architectural and organizational participatory design, while improving staff satisfaction and reducing costs. Jiang et al. [6] proposed an immersive serious game delivered via an HMD for teaching users the sequence of operations necessary to launch a lifeboat and on handling the potential risks of such a task. Furthermore, the analysis of the usage of VR technology in immersive simulations presented by Menin et al. [7] showed that VR enables users to save money and perform safe training,

*Corresponding author:
e-mail: aline.menin@inria.fr (Aline Menin)

while supporting knowledge gain as efficiently as conventional methods. Further, VR training simulations improve knowledge retention and increase users' engagement, which, consequently, enable situational awareness and improve the understanding of training procedures [7].

It is well-known that effective learning occurs when the body senses are stimulated, particularly the visual one [8]. Nonetheless, motion sickness is a common side-effect of using VR due to standing still while the surrounding environment is moving, disturbing the brain's equilibrium and causing nausea. Thus, in this paper, we investigate the effects of different VR display devices and interaction techniques on perceptual learning and knowledge retention. We performed two user-based evaluations with two realistic immersive simulations designed to teach and train people on work safety. We considered aspects of user experience (i.e. workload and cybersickness) and user performance (i.e. error rate, simulation time) to understand how VR technology stimulates the body's senses and how it reacts. The contributions of this work are summarized as follows:

- A user study involving 107 unpaid persons divided into two groups to assess the effects of: three different display devices (from conventional desktop displays to VR headsets) on perceptual learning, and four different interaction techniques (from traditional game like interaction to VR techniques, such as walking-in-place) on perception and knowledge retention. The studies also assessed user experience in two immersive simulations designed for training purposes; and
- A set of lessons learned, which can lead to research questions and help developers on choosing the VR setup for new training simulations.

The remainder of this paper is organized as follows. Section 2 presents the related work. Section 3 explains our design rationale, and presents the description of the serious games developed. Section 4 presents the protocol used for both user studies, while Sections 5 and 6 present the two experiments, the methodology employed, the results achieved and the hypotheses assessment. Section 7 discusses our findings, presents the lessons learned from this study, and the limitations of our studies. Finally, we conclude in Section 8.

2. Related Work

Studies preceding ours have already investigated the impacts of VR technology on user experience (UX), which we present hereafter. Mania et al. [9] have studied the effects of different levels of immersion (real-world, desktop, audio-only, and head-mounted display – HMD) on sense of presence and memorization in a seminar-like presentation. Predictably, users reported the highest sense of presence and performance when they were in a real-world seminar room. In terms of fidelity aspects (i.e. the objective degree of exactness with which real-world sensory stimuli is reproduced [10]), Mania et al. [9] results suggest that low interaction/display fidelity may improve memorization due to the higher attention demand imposed to the cognitive system. These findings were later contradicted by Bowman et al. [11], who showed that a higher level of visual fidelity provides better performance in procedure memorization tasks. They also showed that high display fidelity improves spatial knowledge and allows users to use spatial strategies on procedure memorization, leading to better target detection. More recently, Kwon [12] showed that high fidelity VR setups help users to recognize a virtual experience as the actual experience, while improving their ability to analyze, evaluate, and create questions based on the subject learned during the simulation.

Ragan et al. [13] showed that the use of auxiliary spatial information affects mental strategies and improves user performance in terms of cognitive processing and learning-based activities. Moreover, stereoscopic vision has been shown to improve memory recall when objects are consistent with the environmental context [14]. In another study, Roman et al. [15] showed that a three-monitor CAVE is more engaging with a 3D first-person shooter game than a single monitor.

Napieralski et al. [16] showed that users can locate themselves easier inside a VR environment when the realism of graphics is higher. Furthermore, Ragan et al. [17] showed that high visual realism improves strategy transfer, but worsen user performance in scanning tasks probably due to the extra information that exists on VR environments simulating real-world scenarios (e.g., trees, people, buildings), which can distract the user from their main task. They also demonstrated that training with higher FOVs lead to better object detection in a serious game using scanning tasks to train military personnel. Although, they did not find any correlation between the user performance on the VR environment and the real-world, their findings suggest that a higher FOV do not improve the real-world task performance more than training with a lower FOV.

Bowman and McMahan [18] also showed that using wide FOV and high resolution provides a less cluttered and more comprehensible VE. McMahan et al. [10] showed that high display and interaction fidelity affect strategy and user performance in a first-person shooter game in a way that, with high interaction fidelity users took less damage and were less accurate then with low interaction fidelity. The authors showed that users' familiarity with the technology improves user performance, which could explain the high accuracy when interacting with conventional mouse and keyboard devices. In terms of user experience (UX), the authors observed that high display/interaction fidelity increases the users' sense of presence, engagement and the simulator's assessed usability.

Krokos et al. [19] showed that using virtual memory palaces (i.e. placing pieces of information within palace and medieval town environments, and associating them to salient features of the environment) as a spatial mnemonic to support information recall are more effective when delivered via HMDs than desktop displays. Nabioyuni and Bowman [20] investigated the effects of hyper-natural transfer and biomechanical techniques for navigating inside a VR environment. Their findings suggest that well-designed hyper-natural navigation techniques can be understood and adapted by users, resulting in more speed performance. However, these techniques may still be more difficult to control than real walking when performing complicated,

more precise movements. Further, hyper-natural biomechanical techniques do not improve locomotion performance in VR, but instead reduce accuracy and users' comfort [20].

Kakoschke et al. [21] investigated the effects of Approach-Avoidance Training (AAT) delivered via three interfaces (computer, smartphone, and VR) on user experience (flow, immersion, engagement) and performance (accuracy, approach bias). Their results showed that VR was a more effective way to deliver the training by providing higher engagement, flow, and immersion than the AAT delivered via a computer or smartphone application, while reducing errors.

Albus et al. [22] used signals in the form of textual annotations to support learning in VR simulations. They also investigated its effects on different learning outcomes and cognitive load. Results showed that annotations in VR can help users to process and recall the information, but have no effects in terms of reducing distractions. Gupta et al. [23] used a VR simulator for condylar plating surgery to show that VR based environments can serve both as skills training and learning tools.

Di Mascio et al. [24] evaluated the acceptability, usability, and engagement of two HMDs (i.e. Oculus Rift and Hololens) as tools to provide VR and AR-based treatment for people with Autism Spectrum Disorder (ASD). For both devices, results suggest that initial training is necessary, as well as long-term usage to provide users with freedom to virtually and physically move around the environment. The study showed yet that familiarity with objects in an IVE's (Immersive Virtual Environments) had a higher impact on emotional participation than photo-realism.

Hasanzadeh et al. [25] investigated the feasibility and usefulness of providing passive haptics in a mixed-reality (MR) environment to capture the risk-taking behavior of workers. Results showed that MR helps to raise users' sense of presence and to capture their realistic response to safety features for both research and training purposes. Additionally, the study showed that extroverts workers could place themselves more readily within this mental representation and experience higher levels of involvement, higher degrees of presence, and stronger sense of being there.

Recently, due to the constraints imposed by the COVID-19 pandemics, the use of VR for training and learning has receiving more attention and the influence of aspects such as users' visual attention and behavior are being explored. Simeone et al. [26] found that the presence of a virtual instructor increases the engagement and accelerate the progress of the user during a learning experience. Bozkir et al. [27] focused in three different objects-of-interest for measuring attention: peer-learners, instructor, and lecture material. More specifically, they varied sitting positions of students, visualization styles of virtual avatars (realistic or cartoon-like), and hand-raising percentages of peer-learners. Results showed that such manipulations play an important role in students' attention.

These user studies illustrate the interest and benefits of using VR technologies to provide the user with high fidelity setups to support training procedures through low-cost and danger-free alternatives. A high fidelity VR setup would be built with technologies that provide the closest experience as one would have in the real-world, which means peripheral vision and natural interaction techniques, such as through mid-air gestures and real walking. Nonetheless, we could see that, depending on the target tasks, low interaction fidelity provides better user performance than high fidelity due to the users' familiarity with devices such as game controllers and mouse devices. The study reported by McMahan et al. [10] is the closest to the one we present in this paper in terms of comparing different combinations of low and high interaction/display fidelity devices, which considers for instance the joint use of CAVEs and mouse devices or display-walls and pointing devices. However, the disparity in terms of fidelity between one display and interaction devices are quite large.

In this paper, we extend the understanding of the effects of low/intermediate/high fidelity in user perception and knowledge retention. We considered different VR devices, interaction and navigation techniques on the outcomes of training simulations. Few previous studies have considered the effects of motion sickness, even though it is known to be a side-effect of VR, or the quantity of work necessary to achieve the task using high fidelity interaction tools, which are quite new to the users. Therefore, we measure motion sickness and workload and evaluate their impact on user engagement, sense of presence and immersion.

3. Design Rationale

Perception and learning are intrinsically connected: in training tasks, one cannot learn a new procedure without the ability of perceiving one's surroundings. Thus, this study focuses on investigating the effects of different VR technologies on perception and learning aspects of VR simulations. To reduce cognitive load possibly engendered by the combination of all these technologies, we separated the study in two user experiments: (i) user perception (Section 5) comparing three display devices, and (ii) knowledge gain (Section 6) comparing four combinations of semi- and non-natural interaction and locomotion techniques.

First and foremost, the purpose of a training simulation is to prepare the user for the real-world situation, which has naturally high visual complexity. Thus, the simulators where build on the basis of realistic scenarios, supporting the transfer of what has been learned during the simulation to the real-world situation. We used a simulator for **risk perception assessment** (Fig. 1a), which purpose is to train workers on detecting potential risk elements in a normal workplace environment [28], and a simulator for **lightning rod replacement** (Fig. 1b), which intend to train apprentices on basic safety procedures for electrical installations on public utility poles.

The risk perception assessment simulator reinforces users' perception through their ability to see, hear, or become aware of something through the senses to apprehend their surroundings, detecting and avoiding risk hazards. Thus, the simulator train users through perceptual learning, which comprises the ability to detect pieces of information (i.e. events, distinctive features, and affordances) offered by the environment [29]. Ragan et al. [17] had previously evaluated the effects of different

(a) (b)

Fig. 1: Virtual environments used in the Experiment I - risk perception assessment (a) and Experiment II - lighting rod replacement(b).

levels of FOV (in an HMD device) on a scanning task using realistic scenarios, which shown no significant effect on target detection or assessment strategy usage. In order to extend the understanding of VR impact on perceptual learning, we investigate how different display devices providing different FOVs and different usage settings (more or less comfortable to the user), could impact user experience and performance.

Further to acquire information, the user should be able to retain the information obtained through training simulations in order to transfer it into the real-world situation. Recalling information is also known as knowledge and can be classified into different categories such as factual, conceptual, procedural and meta-cognitive [30]. Both simulators require the user to recognize specific details or elements (i.e. factual knowledge). In this paper, we explore knowledge retention through the lighting rod replacement simulator, where we investigate the effects of different navigation and interaction techniques for grabbing and manipulation, which are fundamental in numerous training simulators (e.g., firefighting, military training, etc). In terms of interaction, we use motion tracking to map knowledge retention to the real-world by providing high fidelity experiences to the user.

In terms of navigation, we considered the use of motion tracking, but the solutions found to track the large physical spaces required for both simulations (Fig. 1) were too expensive [31]. Hyper-natural navigation techniques such as Seven League Boots [32] could improve speed performance, but they might be difficult to control when precision is needed [20]. Thus, we focus in less natural navigation techniques such as walking-in-place (WIP) and joystick navigation. Particularly, the Wii balance board has been largely used by researchers to provide low-cost WIP techniques, showing positive effects on user performance in human-scale spaces and spatial orientation, while providing high sense of presence [33].

In terms of grabbing and manipulating objects, motion tracking can also enable natural interaction through mid-air gestures [34], which has shown to offer less precise control and require more physical work on the part of the user [35]. Furthermore, Mine et al. [36] showed that having a physical reference helps the user to be more precise on memory recall. Thus, we combine physical objects and body movement by using the Razer Hydra controllers, and compare it to conventional game controllers, since users are familiar with these devices.

The first experiment focuses on perception, using a VE that requires walking around while observing and perceiving hazardous features in different rooms. In this task, the display device impacts the performance more than the interaction technique, so we compared three different display devices. The second experiment focused on investigating the effects of interaction and locomotion techniques on learning. Although it uses a different simulator, the perception skills are as important, since the user should be aware of their surroundings to judge the best way of securing the workplace. Therefore, we chose the display device that has shown the best effects on the first trial, ensuring that the display was not a bias, while showing positive effects in terms of perception.

Particularly, as shown by previous works (see Section 2), VR devices providing highly immersive and natural experiences can improve user performance. Therefore, we selected a set of VR devices that provide proper levels of immersion (i.e. HMD and display-wall) and naturalness (i.e. Razer Hydra and Wii Balance Boards) while comparing them with traditional and familiar devices (i.e. PC monitor and game controllers), enabling us to investigate their effects the outcomes of the training simulations. Furthermore, the selected devices are cheap and therefore accessible to the companies interested on using these types of simulators, thus easing their dissemination since these are intended for use in large-scale training sessions within companies [28]. Hereafter, we first describe the shared experiment protocol, measures, and statistical analysis process, then we present both user experiments and their results.

4. Shared Protocol Between Experiments and Measures

Since we focus on learning, we assumed that using a within-subjects design could influence the results by allowing users to recycle the information they learned while using a particular display/interaction technique. Thus, we used a between-subjects design in both experiments, where participants were randomly assigned to each experimental condition.

In both experiments, the experimental sessions were individual, accompanied by an experimenter who guided the participant through five steps of a standard protocol consisting of: (1) a Term and Conditions Agreement, where the user is informed about the risks and benefits of participating in the experiment, and through which they consent their participation allowing us to anonymously use their data for research purposes; (2) a socio-demographic questionnaire, where we gather standard profiling information (e.g., age, sex, profession), data regarding their experience with the technologies used (e.g., experience with VR, 3D games) and the subject being addressed in each simulation (e.g., work safety); (3) a learning phase, where users were given time to get familiar with the VR setup; (4) a trial phase, where the user carry out the required tasks; and (5) a post-test phase, where we apply a questionnaire gathering users' self-reported engagement, cybersickness, and workload, as well as information specific to each experiment.

Particularly, to measure the impact of the VR devices on cybersickness (i.e. visually induced motion sickness) we administered the Simulator Sickness Questionnaire (SSQ) [37]. The questionnaire allows the user to assess the severity with which they experienced sixteen sickness symptoms in a 4-point scale (i.e. from low to high), which allows to determine the user's level of sickness regarding four scales: nausea (N), oculomotor (O), disorientation (D), and total severity (TS). The highest scores possible are 124 for nausea, 90.9 for oculomotor, 97.4 for disorientation, and 108.6 for total severity. As common practice [38], we administrated the SSQ before and after the trial in both experiments to obtain baseline and completion measurements. This way, we can ensure that the results are not biased by the condition of the participant before the trial, who could be stressed, anxious, or nauseous for unrelated reasons.

Regarding statistical analysis, we set the significance level to $p = 0.05$ to analyze the empirical data. We submitted the data to a mixed design, where the between-subjects variables are the experimental conditions and the user profile, while the within-subjects variables are the simulator sickness scores (obtained before and after the trial) and other measures specific to each experiment, such as the sense of presence scores (obtained per room in the VE), and the self-reported workload scores (obtained before and after the trial). We performed a Shapiro-Wilk Normality test of the null hypothesis that the data come from a normal distribution, and a Fligner-Killeen test of the null hypothesis that the variances in each group are the same. If the data passes both tests (p <0.5) we performed a One-Way ANOVA test. Otherwise, the data was submitted to non-parametric tests, namely Friedman for paired and Kruskal-Wallis for unpaired groups. In tests involving more than two groups, with statistical significance, we ran post-hoc tests: Tukey's range test [39] for One-Way ANOVA, and Nemenyi test [40] for the remaining. We report statistical significance in the difference of means for two or more groups through codes on the charts, if any, as follows: '***' 0.001, '**' 0.01, '*' 0.05, and '.' 0.1.

5. Experiment I: User Perception

This experiment aims to understand whether different VR display devices influence the user perception in a simulation designed to train workers of an electricity distribution company on detecting risks elements in their work environment. This experiment requires users to leverage their perception skills to detect potential hazards in the virtual environment. Based on previous studies' findings (see Section 2), we believe that the higher the feeling of being "there" in the VE, the better the user can perform the tasks and recall the information acquired during the simulation. Particularly, a high immersive setup would enable them to completely focus on the elements of the virtual world without being distracted by the real-world objects surrounding them. Thus, we hypothesized that (**H1**) display devices producing higher self-reported sense of presence improve the user performance. Moreover, the existing immersive training simulators have shown overall positive outcomes [7], which lead us to hypothesize that (**H2**) simulator sickness will not worsen the user performance.

5.1. Virtual Environment

This simulation was designed for assessing the ability of workers to detect potential risk elements in a normal workplace environment. The VE contains a building with a reception room, a parking lot, an office, and a kitchen. Further to the normal objects of an office environment, the scenario has 53 objects defined as potential hazards which can be simple or composite. The former refers to imminent risk items (e.g., wet floor, blocked fire extinguisher), while the latter refers to objects which hazard is triggered by an external situation (e.g., flashing lamp, alarm off). These require higher cognitive effort from the user to be detected.

5.2. Experimental Conditions

We compared three VR display devices: (**A**) a 23-inch LCD screen with a 120Hz refresh rate and 1920×1080 pixels resolution **desktop display**, which we used as our baseline condition, since games and current non-immersive simulations are usually ran on PC desktops; (**B**) a 3200×1800 pixels resolution **display wall** built up from a set of twelve 22-inch LCD screens, with a total dimension of 244×108 centimeters (Fig. 2b). This display intends to immerse the user on the simulation while allowing them to be aware of their surroundings; and (**C**) an **HMD** device, since it is widely used in VR and provides high self-reported sense of presence, allowing the user to focus on the simulation while avoiding distractions issued from their surroundings. We used the Oculus Rift DK2 device, which has a 960×1080 pixels resolution per eye and a refresh rate of 75Hz.

5.3. Participants

Sixty-one volunteers (15 female) took part in this experiment. They were students, mostly from Computer Science, and University personnel, aged between 19 and 63 years old (M=28.67, SD=10.12). They were all beginners in terms of work safety and VR, and 42% reported to be experienced on 3D video games. All the participants reported normal or corrected-to-normal vision.

Fig. 2: Experimental conditions for the first experiment (user perception).

5.4. Task

The users were assigned with two tasks: (1) to follow the instructions (e.g. "Inside of this building has a kitchen. Go there and get a coffee.") given by a storyteller, while (2) scanning the environment for potential risks objects. The guidance enabled users to go through every room in the building. In each room, they were instructed to reach a target object (flashing in red) by moving back and forward using the game controller's joystick. Users would turn their heads to set the movement direction when using the HMD and the display wall, and use the joystick when using the desktop display. After reaching the target, the system triggers the next instruction, except in the case of two special instructions, when the user should stand in a location and scan the room for potential risks for as long as they would like. These tasks were included as a strategy to reduce the difference among the experimental conditions, demanding users to perform the same movements by either turning their heads with the HMD or using the game controller. Further, it serves also to verify whether users would simply follow what the narrator was telling them to do or they were also attentive to the scanning task. Such as in the real world, the detection of risky situations should be a background task.

5.5. Measures

We measured the users' subjective sense of presence using the SUS (Slater-Usoh-Steed) presence questionnaire [41], which consists of six statements rated in a 7-point Likert scale used to assess the sense of presence in each room of the VE. Further, we asked users to respond the question "how much immersed did you feel?" in a 10-point scale, in order to assess their sense of immersion. Regarding user experience, we asked users to rate in a 5-point Likert scale the following statements: "I believe the simulation achieve the purpose of assessing the user's aptitude on detecting risk hazards in the work environment", "I could perform the simulation easily"; "I enjoyed using the simulator". Finally, we measured user performance by combining three factors: task completion time (the lower the better), traveled distance in the VE (the less the better), and the number of detected risks (the more the better).

5.6. Results

The virtual path which users were supposed to follow was determined by the audio instructions, which send them to the car at the parking lot, the coffee maker at the kitchen, the desk at the office, and, finally, to meet their co-workers outside the office. Although, we did not observe major differences among the experimental conditions, users wearing the HMD followed a path slightly more straight than the ones in Display-Wall and Desktop conditions.

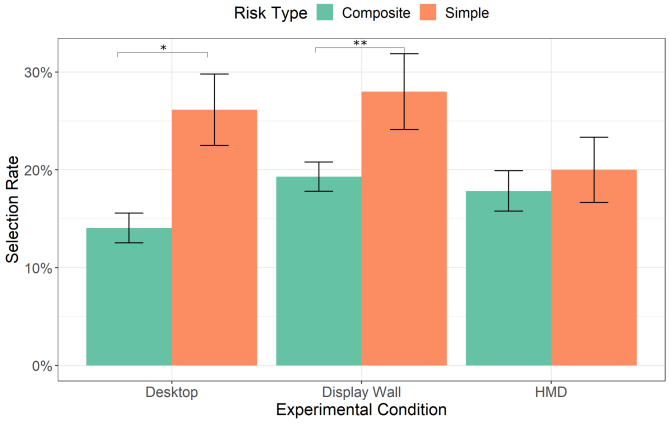

Fig. 3: Mean selection rates per type of risk in each experimental condition of experiment I.

Overall, users took between 5 and 15 minutes to complete the simulation regardless the display device. We observed significant difference in the mean completion times for the display wall and the remaining devices, suggesting that participants using the display wall took longer to finish the simulation (M=12.4, p<0.001). In terms of risk selection, users were able to identify about 21% of potential risk objects in the VE (Fig. 3). The mean scores show that users found more risk objects using the desktop display than with the remaining, but no statistical difference was observed. Nonetheless, when analyzing the type of risk objects users found, we observed that they detected more simple than composite risks using the display wall (p<0.001) and the desktop display (p<0.05), while there was no significant difference under the HMD condition. We observed that completion time and risk selection is positively correlated when using the desktop (p<0.02, R=0.5) and HMD (p<0.03, R=0.48) display devices.

In order to determine whether the devices inflicted cybersickness on users, we compared the SSQ scores measured before and after the trial. We observed statistical difference be-

tween mean scores for symptoms of nausea (p<0.01), disorientation (p<0.001), and total severity (p<0.01) of users using the HMD. The desktop display reduced users' perceived severity of symptoms, where we observed statistical difference in the mean scores of oculomotor and total severity (p<0.02). No statistical difference was observed for the display wall device.

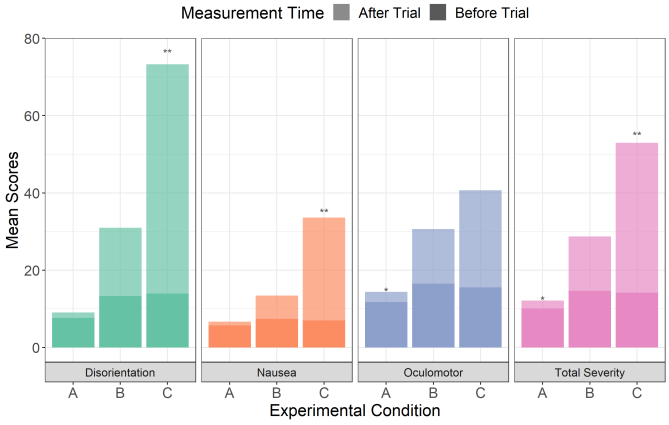

Fig. 4: Mean scores per SSQ scale before and after the trial in each experimental condition of Experiment I. A, B and C are the experimental conditions, see section 5.2

We did not observe statistical differences in the mean SSQ scores for users using the desktop display and the display wall. However, the users of the HMD appear to perceive the severity of symptoms of nausea, disorientation, and the total severity significantly higher than the users of other devices (p<0.01) (Fig. 4). We have also observed statistical differences in the mean scores of the oculomotor scale between the HMD and the desktop display devices (p<0.05), while no statistical differences were observed in regard to the display wall. The highest SSQ score was 45.4 points (SD=53.7) regarding disorientation symptoms while using the HMD device, which is a high score compared to the maximum score of 97.4 points.

In terms of immersion, mean scores suggest that participants using the HMD device (M=7.96, SD=1.68) perceived themselves more immersed in the simulation than the users of the display wall (M=5.8, SD=2.63, p<0.001) and the desktop display (M=5.45, SD=1.77, p<0.01). Regarding the users' sense of presence, we observed statistical significance in the SUS scores for the HMD and the desktop display (p<0.05), indicating that users could feel more present when using the HMD device. No statistical difference was observed in the scores for the display wall and the remaining devices.

Overall, users reported to be satisfied with the experience and, particularly the users who experienced it through the HMD and display wall, reported to have felt "really" there. A few participants mentioned that having to only walk towards what they are seeing did not felt natural. However, we believe that, differently from the real world, looking towards one direction and walking towards another can increase disorientation. Users have also appreciated the short duration of the experiment and agreed that the simulation meets its goals.

5.6.1. *Hypotheses Assessment.*

Although the assessed sense of presence of users using the HMD was higher than the other devices, the participants' risk selection rate and completion time were similar among devices, which means that we cannot accept **H1**. Nonetheless, we observed that the mean number of risk selection was quite low, which is most likely a consequence of the users' overall lack of expertise on work safety. The users with the HMD device reported more severe sickness symptoms than the users of display wall and desktop display devices, which latter appear to have reduced the severity of symptoms. Nonetheless, completion time and risk selection rate did not present any difference among users of different display devices. Moreover, we did not observe correlation between SSQ scores and completion time or risk selection rate, which allows us to accept **H2**. Although the participants did not present high performance, their varied perception of sickness symptoms severity did not impacted the user performance among the different display devices.

6. Experiment II: Knowledge Gain

This experiment intends to explore the effects of two different interaction and two locomotion techniques on user experience and knowledge gain in a simulator designed for teaching apprentices the work safety procedures of a lightning rod replacement task. Overall, we believe the devices enabling movements closer to natural ones would improve user experience and, consequently, increase the knowledge gain from the simulation. Therefore, we hypothesized that (**H1**) the closest the movement is to the natural one, the better will be the user performance; (**H2**) the devices enabling arms and legs movements will not inflict severe motion sickness; and (**H3**) high workload inducing techniques will worsen user performance.

6.1. *Virtual Environment*

This simulation was designed for training apprentices on basic safety procedures for electrical installations on public utility poles. The VE consists of a town street where a lightning rod is being replaced (Fig. 1b). A storyteller guides the user through audio instructions (e.g. "The workers need to be equipped with scrap gloves, safety belt, and helmet, while carrying the service order. Please, select the objects representing this equipment. They are flashing in red. When you have finished, activate the next instruction."), while teaching them the steps needed to secure the job site for workers and pedestrians. The simulation has three phases: a learning phase, where the user learn the proper placement of safety items, which is indicated via glowing objects and arrows; and two application phases, where the user is asked to perform the safety task using the information they just learned. In the learning phase, users can only grab objects that have been mentioned by the storyteller, which should reinforce their focus on the task. In the application phases, users can grab whichever items they like and follow their own strategy to place them in the scene. We change the initial placement of objects between one phase and another to reduce the bias of already knowing where to start.

The users were asked to perform a set of five tasks: (1) to ensure that the truck is parked correctly and stuck with shims at the wheels; (2) to equip workers with scrap gloves, safety belt, helmet, and the service order; (3) to place the materials on the protection canvas; (4) to signal the workplace using signposts; and (5) to isolate the workplace, including the vehicle, using cones and ropes. The tasks should be performed according to the users' interpretation. In the assignment to arrange the cones around the workplace, the users had eight available cones to be arranged along a square illustrated by four arrows, so they could either place one cone by each arrow or distribute a number of cones they found necessary in the provided space. Later, they were requested to tie the ropes between the cones previously placed. In this case, there were seven ropes available and they could choose a number of ropes they found necessary to isolate the place. A barrier prevented users from crossing through the rope after it has been tied between two cones, which served to remind them of leaving an opening in the isolated area to allow workers to transit.

6.2. Experimental Conditions and Apparatus

In this experiment we compared two interaction and two locomotion techniques. For grabbing and manipulating objects, we use (1) a **ray-casting** technique based on the Razer Hydra controllers used as pointing devices controlled by the user arms' movements. When the object is within view, the user holds the controller's button to grab and drag it, and drop it by releasing the button; and (2) a **head orientation** technique, which uses the user's head orientation to point out to objects and one of the Hydra's controllers to manipulate them using a button. In terms of locomotion techniques, we use (1) the **joystick** of a Hydra's controller for moving forward, backward, left, and right in the environment; and (2) a **WIP** technique supported by a set of four Wii balance boards, which gives the user a 0.80×1.2 meters platform to "walk". The **experimental conditions** for this experiment are then resulting from the combination of these techniques: (**A**) ray-casting and WIP, (**B**) ray-casting and joystick, (**C**) head orientation and WIP, and (**D**) head orientation and joystick.

The equipment worn by the participants consisted of an Oculus Rift DK2 device, a headphone, and a Sixense Razer Hydra. To ensure a certain homogeneity among the conditions, users went through the simulation on top of the Wii Balance Boards platform and used the Razer Hydra controllers, regardless of the enabled interaction/locomotion techniques (Fig. 5). The Razer Hydra device contains a base station with a low-power magnetic field that gets the controller's position and orientation. Although it does not require a line of sight to the controllers, the range area is quite small, introducing noise to the tracking when the user turns its back to the base station. Therefore, we used a band to hold the base station on the user's body (Fig. 5b). In between, we placed a smartphone to capture the orientation of the user's torso, allowing this way to properly place the virtual hands according to the user's rotational movements. Finally, to provide users with confidence and control when using the WIP technique, we added an EVA mat over the platform with EVA borders around 7.5 centimeters high so the user could sense it, being aware of the platform limits.

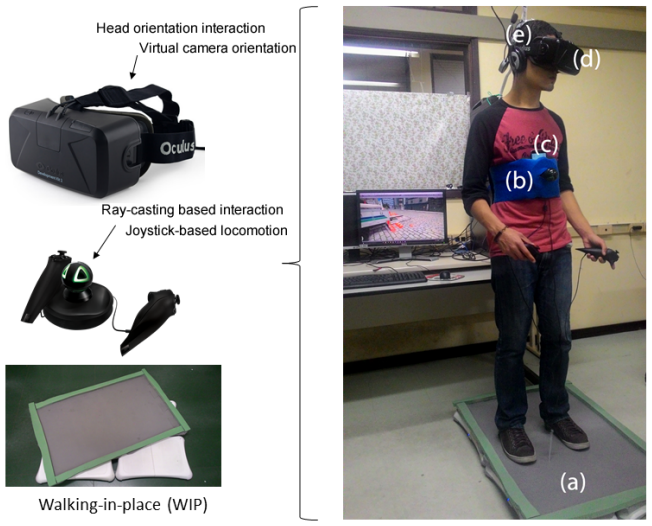

Fig. 5: Technical apparatus of experiment II: (a) WIP platform; (b) band to hold the Razer Hydra base station on the user's body; (c) a smartphone to get the user's orientation; (d) Oculus Rift device; (e) headphones.

6.3. Participants

Sixty volunteers took part in this experiment. They were all students recruited in the context of a Human-Computer Interaction course. Thirteen participants abandoned the simulation due to sever cybersickness symptoms, and we have lost the log files of one participant. Thus, we conducted the experiment with a sample of 46 participants (4 female), aged between 18 and 33 years old (M=23.21, SD=3.42). They reported to have great experience with 3D video games, but little with VR. They all had normal or corrected-to-normal vision (self-reported).

6.4. Measures

Prior to the trial, we established the baseline for cybersickness and users' knowledge regarding safety procedures. We also assessed users' subjective memory complaints through a 10-item memory questionnaire [42], which information could support knowledge outcomes. To assess knowledge (before and after the trial), we prepared a 7-question test on the safety procedures addressed in the simulation. These were essay questions, presented one by one to the user and defined as follows: *What do you do to secure a job site for workers and pedestrians?*; *What do you do after arriving at the site and parking the vehicle?*; *What safety equipment do you wear to perform the task?*; *What extra item(s) do you need?*; *Where and how do you place the working tools?*; *What do you do first to secure the site for pedestrians?*; *How do you isolate the job site?*.

We applied the NASA Task Load Index (TLX) [43] questionnaire to assess workload for using the different navigation and interaction techniques. The questionnaire provides an overall workload score based on a weighted average of ratings on six workload factors (scales): Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Own Performance (OP), Effort (EF), and Frustration (FR). Additionally, we assess user engagement's factors such as immersion, presence, flow, and

physiological absorption through the Game Engagement Questionnaire (GEQ) [10]. The GEQ consists of 15 statements describing different feelings, which the user must rate in a 5-point Likert scale to indicate the intensity with which they experienced that feeling during the simulation. Finally, we measure user performance by combining three information: task completion time, accuracy of objects' placement according to the instructions given in the learning phase, and knowledge gain.

6.5. Results

Fig. 6 presents the mean scores per SSQ scale and experimental condition. Although we have found statistical difference in the mean scores of certain scales before and after the trial in every experimental condition, suggesting that they inflict cybersickness to users, we did not observe statistical difference in the mean scores among the experimental conditions, neither before nor after the trial. Particularly, every experimental condition has engendered disorientation, which highest scores (M=56.84) were reported by users of condition C (head orientation and WIP). Surprisingly, we found a 82% correlation between disorientation and knowledge gain in condition A (ray-casting and WIP) with p<0.01. Although no statistical difference was observed, the users in this group reported the lowest disorientation scores and the highest experience scores in 3D games (M=4.25 in a 5-point scale, SD=1.03), suggesting that disorientation probably did not affect their performance.

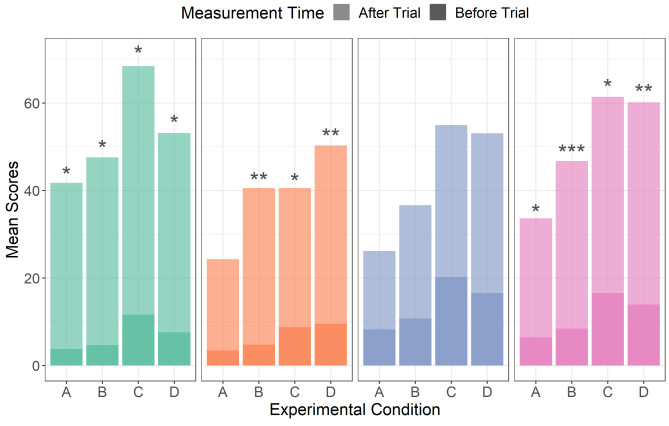

Fig. 6: Mean scores per SSQ scale before and after the trial in each experimental condition of Experiment II. A, B, C and D are the experimental conditions, see section 6.2

The overall workload scores were above average (scores range from 0 to 100), varying from 49.2 in condition A to 62.4 in condition B, which difference of means showed statistical significance with p<0.05. In terms of the different scales measured by the NASA TLX questionnaire (Fig. 7), we observed statistical difference in the mean scores of: frustration between conditions A (M=37.27) and D (M=206.36); mental demand between conditions A (M=177.27) and D (M=50); own performance between conditions B (M=315) and D (M=143.63); and physical demand between conditions C (M=131.66) and D (M=205.45). Furthermore, we observed a 71% correlation between mental demand and knowledge gain in condition A (p<0.05), suggesting that users had higher knowledge gain

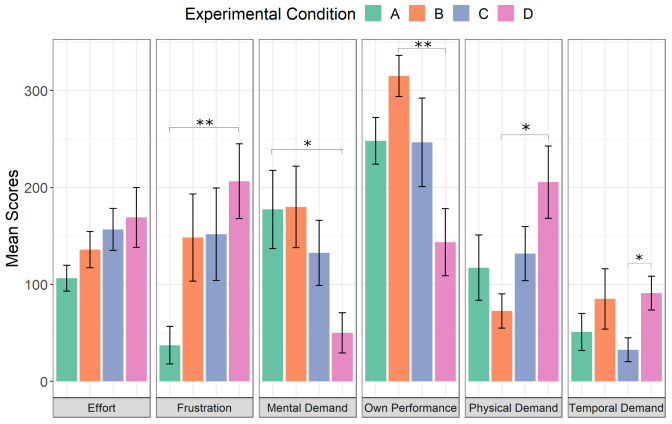

Fig. 7: Reported workload scores for each scale: mental demand (MD), physical demand (PD), temporal demand (TD), own performance (OP), effort (EF) and frustration (FR) for the experimental conditions on Experiment II.

when the task required higher mental effort. For users in condition B, we found a -69% correlation between overall workload scores and knowledge gain (p=0.019), suggesting that when users needed higher loads of work their ability to retain knowledge is reduced.

Figure 8 shows the response accuracy for the knowledge test taken by users before, after, and three weeks after the trial. We observed a statistical difference in the mean accuracy scores for those three measurement times. We did not found statistical difference between scores of tests taken after the trial, which similarity between means suggest that users were capable of recalling the information learned during the simulation. The mean simulation times across conditions varied from 20 minutes in conditions A and B to 32.7 minutes in condition D, which difference of means showed statistical significance (p<0.001). Overall, the completion time reduced between the simulation's phases, which was expected due to users familiarizing with the navigation/interaction techniques, the VE, and the tasks. Users took around 3.6 minutes less time to finish the first application phase than the learning, and the completion time of the second application phase was yet reduced in around 1.1 minutes (no statistical difference was observed among conditions).

Overall, users could complete the simulation's tasks with 95% of accuracy. Particularly, we noticed that 88% of users could correctly recall the instructions given in the learning phase regarding the objects' placement. In terms of isolating the job site, i.e. by forming a fence with cones and a rope around the area containing the canvas with the tools and the truck, while leaving an opening on it to allow workers to transit, we observed that the 78% of users placed the cones correctly, but only 36% of them left an opening in the fence. We observed statistical difference among conditions in users' accuracy to complete certain tasks. In terms of tools placement on the canvas, participants in condition B performed better than users in condition D (p<0.038). Participants in condition C and D were more accurate in terms of arranging the signposts than users in condition A (p<0.05) and B (p<0.05), respectively. Users in condition C were also more accurate in terms of isolating the job site than participants in condition D (p<0.05).

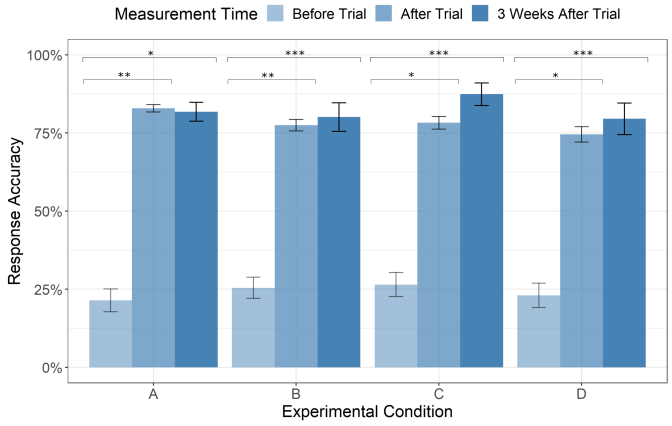

Fig. 8: Response accuracy for the knowledge test taken by users before, after, and three weeks after the trial phase of experiment II.

The GEQ mean score was 2.8 across experimental conditions, which differences were not statistical significant. We have found a 45% correlation between mental demand and engagement (p<0.01), suggesting that navigation/interaction techniques demanding higher mental effort also tend to be more engaging. Overall, users reported to have enjoyed the experience and found it easy to learn and apply the safety procedures. As in the previous experiment, we allow users to walk towards their looking direction to reduce the severity of sickness symptoms. In this matter, two persons have said that they would like to walk backwards. Selecting objects in the scene was reported to be difficult by some users, particularly regarding the cones. Participants suggested the addition of a checklist to allow the user to check tasks as done as they progress in the simulation.

6.6. Hypotheses Assessment

Due to the lack of statistical significance regarding knowledge gain, simulation time, and task accuracy across conditions, we cannot accept not refuse **H1**. Every experimental condition increased the severity of sickness symptoms in every SSQ scale, except oculomotor. Condition A requires the user to move their arms around to select objects and employs the WIP technique for navigating in the VE. Although the mean SSQ scores in condition A are the lowest in every scale, we could not find statistical significance to support the difference of means among conditions. In terms of knowledge gain, a negative correlation between overall workload and knowledge gain suggests that high demanding navigation/interaction techniques could distract users from the task, requiring them to place higher attention on how to use the techniques, decreasing knowledge gain. In condition A, we observed a positive correlation between mental demand and knowledge gain, suggesting that mentally demanding techniques do not affect knowledge gain. In this condition, we also observed that users appear to be more engaging when mental demand increase. Thus, we believe that the freedom provided by more natural movements in condition A allowed users to keep focus on the task, increasing mental demand and learning. These experiments should be repeated and the sample enlarged to evaluate **H2** and **H3**. However, we believe these results favor their acceptance.

7. Discussion

In this work, we performed two empirical studies to investigate the impacts of three different VR displays and four interaction techniques on several VR aspects, such as motion sickness, workload, engagement, sense of presence, user performance, and learning outcomes on training simulations.

7.1. Main Findings

Our results showed that participants could perform certain tasks with higher accuracy, such as isolating the job site using the conditions where navigation was supported via the WIP technique, and placing tools over a canvas in conditions where they would grab and manipulate objects using the ray-casting technique. The latter provided a physical reference through a Razer Hydra's controller, which higher accuracy on performing the task suggests that users could better recall the information given during the learning phase, in conformity with previous findings [36]. Although we did not observe statistical difference in knowledge gain among experimental conditions and replication of these experiments may be necessary to evaluate our hypotheses, our results suggest that semi-natural navigation/interaction techniques could improve user performance.

Previous studies suggest that semi-natural interfaces, i.e. interfaces that use real walking and body interaction, could be inefficient since they require the user to operate it differently from what they are used to, which adds workload to the process [44]. However, we observed that users reported higher frustration scores when using the head orientation technique and the joystick to navigate, while they were more engaged with the simulation when moving their arms and legs through the ray-casting and WIP techniques, which were considered high mentally demanding by users. Users also perceived a higher quality on their own performance when using the WIP technique compared to the joystick. Furthermore, we observed a positive correlation between mental demand and knowledge gain in the condition using WIP and ray-casting, which also had the lowest simulator sickness effects. Our results suggest that providing close to natural interaction interfaces can be enjoyable and aid users to transfer the potential mental burden generated from adapting their real-walking expertise to better performance.

7.2. Lessons Learned

Disorientation is probably a consequence of the user unawareness of their surroundings. As we have hypothesized, techniques providing semi-natural interaction did not increase the severity of sickness symptoms, because users could "move" together with the VE. Nonetheless, we observed that disorientation would be a constant across the experimental conditions in experiment II, which coincides with the disorientation reported by users of the HMD in experiment I, suggesting that disorientation is only the fact that the user is not aware of their surroundings, which seems not to be affected by the different interaction/locomotion techniques. Thus, our findings suggest that better VEs would allow the user to be aware of their surroundings, such as through the use of physical interaction devices and more importantly through haptic interaction, allowing the

user to physically feel what they're virtually experiencing. Such physical approaches have been previously explored to increase awareness in VEs and improve user experience (see [45, 46]).

The use of interaction techniques as natural as possible continue to be the best choice. Simulations designed for VR technology aim to provide a real-world-like experience, where high sense of presence, immersion and realism are essential ingredients. Thus, it is important to allow the use of our well-known movements to navigate and interact with the VE. While much research is still needed to improve this technology, there are several studies [10, 20, 32], including this one, showing the advantages of natural or semi-natural interaction techniques on user performance and experience. Thus, users could leverage of training simulators employing well-drafted (semi-)natural interaction techniques, which provide the sense of naturally navigating and grabbing objects in the VE, while providing accuracy for effectively dealing with small objects or performing precise tasks, such as painting.

Perception is more likely to be affected by the user expertise on the simulator subject than by the technology employed. We have observed in experiment I (Section 5) that neither the VR display nor the sickness caused by it affected user perception. However, we noticed that users had a rather poor performance on what refers to the risk scanning task. Users were able to select 20% of risks elements (consistently throughout the devices), which is a low selection rate, likely explained by users' self-reported very low knowledge on work safety procedures. Previous works showed that the familiarity of users with the objects in the environment improves the user experience [24, 10]. Similarly, considering that our simulator have been developed to assess the ability of users to detect hazardous objects in the workplace, we suppose that the knowledge on the subject could have strongly impacted user performance, since one should know what is an hazardous element in the real workplace to be able to identify it in the simulation. Therefore, further studies using the "risk perception assessment" or similar simulations to investigate user perception should recruit target users of the simulation or to provide a training on work safety to the participants before undertaking the simulation.

Gaming experience plays an important role on VR simulation. We have found a positive correlation between disorientation and knowledge gain in condition A of experiment II (Section 6), which oddly suggests that users improved their scores despite the severity of disorientation's symptoms. However, we noticed that the users in this group reported to have high experience with 3D video games. Thus, we could suppose that users' familiarity with games and the constantly experienced mild cybersickness due 3D graphics allowed them to keep focused on the simulation despite the disorientation experienced in the VE. Further, while studying the impact of narrative in immersive VR games, Weech et al. [47] also observed differences cybersickness symptoms reported by gamers and non-gamers. While non-gamers reported higher cybersickness when receiving minimal narrative, regular gamers would present similar levels of cybersickness regardless of the narrative context.

Although our case study focuses on work safety training, our results and lessons learned can assist the choice of VR devices to provide a comfortable and efficient virtual training environment in other application domains and, particularly, for simulations requiring the users to leverage their perception skills. For instance, our observations regarding the effects of users' profile on perception and cybersickness could help designers to conceive suitable guidance for simulators based on users' profile (e.g., whether they need more pre-training on the simulator's subject or the usage of devices, etc). Likewise, the usage of high fidelity devices to provide a real-world-like experience is independent of the application domain and can be generalized to any type of training simulator.

7.3. Limitations

Each user experiment tested multiple independent variables by 107 volunteers in total. Nevertheless, we applied between-subjects designs in both experiments due to the learning nature of our dependent variables, reducing the groups to 20 and 12 persons, respectively. Thus, these experiments should be replicated using larger groups, and preferably the simulator's target users, to strengthen our findings. The lack of expertise on work safety may have introduced bias to the results. As suggested by Nazir et al. [48], immersive simulators are not effective per se and need guidance of a professional in the field when the task is unfamiliar. We have used VR technology that was available to us at the time of this study. Hence, a new set of experiments should be performed using novel technology to understand whether these results are technology independent.

8. Conclusion

We conducted two empirical studies to investigate the effects of VR interface on perceptual learning and knowledge retention in simulations designed for worker safety training. We measured simulator sickness, self-reported workload, engagement, presence and performance. Our objective was to better understand the impacts of different display devices and interaction techniques on user performance and simulation outcomes. Our results strongly suggest that a setup employing interaction and navigation techniques closest to natural movements can improve user performance. In our experiments, we observed this improvement through accurately placing objects in the VE, giving users the feeling of performing better, and having low effects on simulator sickness. However, users still had a physical reference for moving the hands in the VE and select objects, which is supported by previous studies as a mechanism to improve memory recall. Different VR display devices did not influence the user performance in our risk scanning task, but rather their experience on work safety.

9. Acknowledgments

This study was financed in part by the Coordination for the Improvement of Higher Education Personnel, Brazil (CAPES, Finance Code 001). We also acknowledge the support from National Council for Scientific and Technological, Brazil (CNPq), the participants of our experiments for allowing us to borrow their time and knowledge for this study, and the comments of the reviewers that helped us to improve the quality of this paper.

References

[1] Chittaro, L, Buttussi, F. Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. IEEE Transactions on Visualization & Computer Graphics 2015;21(4):529–538. doi:10.1109/TVCG.2015.2391853.

[2] Backlund, P, Engström, H, Hammar, C, Johannesson, M, Lebram, M. Sidh-a game based firefighter training simulation. In: Information Visualization, 2007. IV'07. 11th International Conference. IEEE; 2007, p. 899–907.

[3] Pedraza-Hueso, M, Martín-Calzón, S, Díaz-Pernas, FJ, Martínez-Zarzuela, M. Rehabilitation using kinect-based games and virtual reality. Procedia Computer Science 2015;75:161 – 168. doi:https://doi.org/10.1016/j.procs.2015.12.233; 2015 International Conference Virtual and Augmented Reality in Education.

[4] Cecil, J, Gupta, A, Pirela-Cruz, M. An advanced simulator for orthopedic surgical training. International journal of computer assisted radiology and surgery 2018;13(2):305–319.

[5] Mobach, MP. Do virtual worlds create better real worlds? Virtual Reality 2008;12(3):163–179. URL: https://doi.org/10.1007/s10055-008-0081-2. doi:10.1007/s10055-008-0081-2.

[6] Jiang, M, Lan, W, Chang, J, Dodwell, M, Jekins, J, Yang, HJ, et al. A game prototype for understanding the safety issues of a lifeboat launch. Virtual Real 2018;22(2):137–148. URL: https://doi.org/10.1007/s10055-018-0334-7. doi:10.1007/s10055-018-0334-7.

[7] Menin, A, Torchelsen, R, Nedel, L. An analysis of vr technology used in immersive simulations with a serious game perspective. IEEE Computer Graphics and Applications 2018;38(2):57–73. doi:10.1109/MCG.2018.021951633.

[8] Laird, D. Approaches to training and development addison-wesley. Reading, Mass 1985;.

[9] Mania, K, Chalmers, A. The effects of levels of immersion on memory and presence in virtual environments: a reality centered approach. Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society 2001;4(2):247—264. URL: https://doi.org/10.1089/109493101300117938. doi:10.1089/109493101300117938.

[10] McMahan, RP, Bowman, DA, Zielinski, DJ, Brady, RB. Evaluating display fidelity and interaction fidelity in a virtual reality game. IEEE Transactions on Visualization and Computer Graphics 2012;18(4):626–633. doi:10.1109/TVCG.2012.43.

[11] Bowman, DA, Sowndararajan, A, Ragan, ED, Kopper, R. Higher levels of immersion improve procedure memorization performance. In: Proceedings of the 15th Joint Virtual Reality Eurographics Conference on Virtual Environments. JVRC'09; Aire-la-Ville, Switzerland, Switzerland: Eurographics Association. ISBN 978-3-905674-20-0; 2009, p. 121–128. URL: http://dx.doi.org/10.2312/EGVE/JVRC09/121-128. doi:10.2312/EGVE/JVRC09/121-128.

[12] Kwon, C. Verification of the possibility and effectiveness of experiential learning using hmd-based immersive vr technologies. Virtual Reality 2018;URL: https://doi.org/10.1007/s10055-018-0364-1. doi:10.1007/s10055-018-0364-1.

[13] Ragan, ED. Supporting learning through spatial information presentations in virtual environments. Ph.D. thesis; Blacksburg, VA, USA; 2013. AAI3585819.

[14] Bennett, A, Coxon, M, Mania, K. The effect of stereo and context on memory and awareness states in immersive virtual environments. In: Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization. ACM; 2010, p. 135–140.

[15] Roman, F, Maciel, A, Nedel, L. Improving gameplay in first person 3-d games using multiple displays. Comput Entertain 2015;12(2):1:1–1:22. URL: http://doi.acm.org/10.1145/2701657.2701653. doi:10.1145/2701657.2701653.

[16] Napieralski, PE, Altenhoff, BM, Bertrand, JW, Long, LO, Babu, SV, Pagano, CC, et al. Effects of immersion on spatial updating in virtual panoramas. In: Proceedings of the ACM Symposium on Applied Perception. SAP '12; New York, NY, USA: ACM. ISBN 978-1-4503-1431-2; 2012, p. 129–129. URL: http://doi.acm.org/10.1145/2338676.2338711. doi:10.1145/2338676.2338711.

[17] Ragan, ED, Bowman, DA, Kopper, R, Stinson, C, Scerbo, S, McMahan, RP. Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. IEEE Transactions on Visualization and Computer Graphics 2015;21(7):794–807. doi:10.1109/TVCG.2015.2403312.

[18] McMahan, RP, Bowman, DA. Virtual reality: How much immersion is enough? Computer 2007;40:36–43. URL: doi.ieeecomputersociety.org/10.1109/MC.2007.257. doi:10.1109/MC.2007.257.

[19] Krokos, E, Plaisant, C, Varshney, A. Virtual memory palaces: immersion aids recall. Virtual Reality 2018;URL: https://doi.org/10.1007/s10055-018-0346-3. doi:10.1007/s10055-018-0346-3.

[20] Nabioyuni, M, Bowman, DA. An evaluation of the effects of hyper-natural components of interaction fidelity on locomotion performance in virtual reality. In: Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments. ICAT - EGVE '15; Aire-la-Ville, Switzerland, Switzerland: Eurographics Association. ISBN 978-3-905674-84-2; 2015, p. 167–174. URL: http://dx.doi.org/10.2312/egve.20151325. doi:10.2312/egve.20151325.

[21] Kakoschke, N, Page, R, de Courten, B, Verdejo-Garcia, A, McCormack, J. Brain training with the body in mind: Towards gamified approach-avoidance training using virtual reality. International Journal of Human-Computer Studies 2021;151:102626.

[22] Albus, P, Vogt, A, Seufert, T. Signaling in virtual reality influences learning outcome and cognitive load. Computers & Education 2021;166:104154.

[23] Gupta, A, Cecil, J, Pirela-Cruz, M. A cyber-human based integrated assessment approach for orthopedic surgical training. In: 2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH). IEEE; 2020, p. 1–8.

[24] Di Mascio, T, Tarantino, L, De Gasperis, G, Pino, C. Immersive virtual environments: a comparison of mixed reality and virtual reality headsets for asd treatment. In: International Conference in Methodologies and intelligent Systems for Techhnology Enhanced Learning. Springer; 2019, p. 153–163.

[25] Hasanzadeh, S, Polys, NF, de la Garza, JM. Presence, mixed reality, and risk-taking behavior: A study in safety interventions. IEEE Transactions on Visualization and Computer Graphics 2020;26(5):2115–2125. doi:10.1109/TVCG.2020.2973055.

[26] Simeone, AL, Speicher, M, Molnar, A, Wilde, A, Daiber, F. LIVE: the human role in learning in immersive virtual environments. In: Borst, CW, Kulshreshth, AK, Bruder, G, Serafin, S, Sandor, C, Johnsen, K, et al., editors. Symposium on Spatial User Interaction, SUI 2019, New Orleans, LA, USA, October 19-20, 2019. ACM; 2019, p. 5:1–5:11. URL: https://doi.org/10.1145/3357251.3357590. doi:10.1145/3357251.3357590.

[27] Bozkir, E, Stark, P, Gao, H, Hasenbein, L, Hahn, J, Kasneci, E, et al. Exploiting object-of-interest information to understand attention in VR classrooms. In: IEEE Virtual Reality and 3D User Interfaces, VR 2021, Lisbon, Portugal, March 27 - April 1, 2021. IEEE; 2021, p. 597–605. URL: https://doi.org/10.1109/VR50410.2021.00085. doi:10.1109/VR50410.2021.00085.

[28] Nedel, L, de Souza, VC, Menin, A, Sebben, L, Oliveira, J, Faria, F, et al. Using immersive virtual reality to reduce work accidents in developing countries. IEEE computer graphics and applications 2016;36(2):36–46.

[29] Adolph, KE, Kretch, KS. Gibson's theory of perceptual learning. i: Keller H(Developmental Section red) International Encyclopedia of the Social and Behavioral Sciences, 2015;10:127–134.

[30] Krathwohl, DR. A revision of bloom's taxonomy: An overview. Theory Into Practice 2002;41(4):212–218. doi:10.1207/s15430421tip4104_2.

[31] Hodgson, E, Bachmann, E, Waller, D. Redirected walking to explore virtual environments: Assessing the potential for spatial interference. ACM Trans Appl Percept 2008;8(4):22:1–22:22. URL: http://doi.acm.org/10.1145/2043603.2043604. doi:10.1145/2043603.2043604.

[32] Interrante, V, Ries, B, Anderson, L. Seven league boots: A new metaphor for augmented locomotion through moderately large scale immersive virtual environments. In: 2007 IEEE Symposium on 3D User interfaces. IEEE; 2007,.

[33] Usoh, M, Arthur, K, Whitton, MC, Bastos, R, Steed, A, Slater, M, et al. Walking > walking-in-place > flying, in virtual environments. In: Proceedings of the 26th Annual Conference on Computer Graph-

ics and Interactive Techniques. SIGGRAPH '99; New York, NY, USA: ACM Press/Addison-Wesley Publishing Co. ISBN 0-201-48560-5; 1999, p. 359–364. URL: http://dx.doi.org/10.1145/311535.311589. doi:10.1145/311535.311589.

[34] Greunke, L, Sadagic, A. Taking immersive vr leap in training of landing signal officers. IEEE transactions on visualization and computer graphics 2016;22(4):1482–1491.

[35] Bowman, DA, Hodges, LF. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In: Proceedings of the Symposium on Interactive 3D Graphics. Providence, RI, USA; 1997, p. 35–38.

[36] Mine, MR, Brooks Jr, FP, Sequin, CH. Moving objects in space: exploiting proprioception in virtual-environment interaction. In: Proceedings of the 24th annual conference on Computer graphics and interactive techniques. 1997, p. 19–26.

[37] Kennedy, RS, Lane, NE, Berbaum, KS, Lilienthal, MG. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. The international journal of aviation psychology 1993;3(3):203–220.

[38] Balk, SA, Bertola, MA, Inman, VW. Simulator sickness questionnaire: Twenty years later. In: Seventh International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design. University of Iowa; 2013,.

[39] Abdi, H, Williams, LJ. Tukey's honestly significant difference (hsd) test. Encyclopedia of research design 2010;3(1):1–5.

[40] Nemenyi, PB. Distribution-free multiple comparisons. Princeton University; 1963.

[41] Usoh, M, Catena, E, Arman, S, Slater, M. Using presence questionnaires in reality. Presence 2000;9(5):497–503. doi:10.1162/105474600566989.

[42] Zelinski, EM, Gilewski, M. A 10-item rasch modeled memory self-efficacy scale. Aging & mental health 2004;8(4):293–306.

[43] Hart, SG, Staveland, LE. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In: Hancock, PA, Meshkati, N, editors. Human Mental Workload; vol. 52 of *Advances in Psychology*. North-Holland; 1988, p. 139 – 183. doi:https://doi.org/10.1016/S0166-4115(08)62386-9.

[44] Nabiyouni, M, Saktheeswaran, A, Bowman, DA, Karanth, A. Comparing the performance of natural, semi-natural, and non-natural locomotion techniques in virtual reality. In: 2015 IEEE Symposium on 3D User Interfaces (3DUI). 2015, p. 3–10. doi:10.1109/3DUI.2015.7131717.

[45] de Jesus Oliveira, VA, Brayda, L, Nedel, L, Maciel, A. Designing a vibrotactile head-mounted display for spatial awareness in 3d spaces. IEEE transactions on visualization and computer graphics 2017;23(4):1409–1417.

[46] Valkov, D, Linsen, L. Vibro-tactile feedback for real-world awareness in immersive virtual environments. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE; 2019, p. 340–349.

[47] Weech, S, Kenny, S, Lenizky, M, Barnett-Cowan, M. Narrative and gaming experience interact to affect presence and cybersickness in virtual reality. International Journal of Human-Computer Studies 2020;138:102398.

[48] Nazir, S, Kluge, A, Manca, D. Can immersive virtual environments make the difference in training industrial operators. Proceedings of the Human Factors and Ergonomics Society Europe 2014;:251–265.