Evaluating the Effects of Virtual Reality Environment Learning on Subsequent Robot Teleoperation in an Unfamiliar Building

Karl Eisenträger, Judith Haubner, Jennifer Brade, Wolfgang Einhäuser, Alexandra Bendixen, Sven Winkler, Philipp Klimant, and Georg Jahn

Abstract— Using a map in an unfamiliar environment requires identifying correspondences between elements of the map's allocentric representation and elements in egocentric views. Aligning the map with the environment can be challenging. Virtual reality (VR) allows learning about unfamiliar environments in a sequence of egocentric views that correspond closely to the perspectives and views that are experienced in the actual environment. We compared three methods to prepare for localization and navigation tasks performed by teleoperating a robot in an office building: studying a floor plan of the building and two forms of VR exploration. One group of participants studied a building plan, a second group explored a faithful VR reconstruction of the building from a normal-sized avatar's perspective, and a third group explored the VR from a giant-sized avatar's perspective. All methods contained marked checkpoints. The subsequent tasks were identical for all groups. The self-localization task required indication of the approximate location of the robot in the environment. The navigation task required navigation between checkpoints. Participants took less time to learn with the giant VR perspective and with the floorplan than with the normal VR perspective. Both VR learning methods significantly outperformed the floorplan in the orientation task. Navigation was performed quicker after learning in the giant perspective compared to the normal perspective and the building plan. We conclude that the normal perspective and especially the giant perspective in VR are viable options for preparing for teleoperation in unfamiliar environments when a virtual model of the environment is available.

Index Terms— Human computer interaction (HCI), human-centered computing, virtual reality, human factors, teleoperation, robot, virtual reality

1 Introduction

Virtual reality (VR) has seen a stellar rise in recent years due to advances in technology and its affordability [1]. While VR has turned into a consumer product, it is at the same time deployed to an increasing variety of industries, and research on – and with – VR has continuously been increasing. The malleability of VR renders it a useful tool in entertainment, industry, education, and research. Owing to its adaptability with regards to scale and richness in detail and to the possibility to form abstractions of the object of interest, VR has been used to visually explore many kinds of environments and abstract objects, such as subatomic particles or time-series data [2–4].

1.1 Research gap: VR for subsequent teleoperation

Abstractions and metaphors can be employed in VR to visualize everyday topics. For navigating actual spaces, technological advancement has largely replaced paper maps by digital maps. Such digital maps often feature wayfinding algorithms, can be changed in scale and may include additional information about the area of interest [5]. Moreover, they enable zooming and blending visualization techniques, which allow the user to virtually "visit" the whole world in VR [6]. This leads us to expect that the use of VR as a tool to aid everyday navigation and orientation will increase. This raises the question whether studying an unfamiliar environment through VR improves navigation and orientation compared to studying a classical map, especially if neither map nor VR are accessed while actually operating in the environment. In the present study, we tested whether training with VR better supports subsequent tele-controlled navigation through a building floor and orientation in it compared to the use of a map, and whether there are differences between different VR approaches. Although training navigation through specific environments is highly relevant for teleoperation (cf. section 1.2) and the role of various VR perspectives has been tested in other contexts (sections 2.2, 2.3), including *during* teleoperation, the critical direct comparison to the "vintage" approach of map usage has so far been lacking.

1.2 Challenges of teleoperation

Navigating through and orienting in unfamiliar environments can be a particular challenge when users are not physically present themselves, but instead remotely operate a robot from a distance. The use of such "telerobots" is favored in certain situations, in particular if the operating environment is dangerous or impractical [7, 8]. Impractical in that context can mean inaccessibility for humans (e.g., outer space, deep sea, radiation-contaminated areas), lack of economic viability, or harmful environmental impact (e.g., requiring long-distance travel of specialists). Ranging from the use in the COVID-19 pandemic [9], search and rescue missions, the use in terrestrial mining, in the military, and in space, many use cases for telerobots have been described [10]. The remote control of robots requires spatial knowledge about the environment that the robot is moved through and knowledge about the physicality of the robot that is being controlled [11]. The quality of the connection and in particular its latency, the richness of the (visual) feedback, the familiarity of the operator with the environment as well as with the robot all influence the ability to orient and navigate the robot [12, 13]. The potential benefits of teleoperation (increased safety, reduced travel time and expenses [9, 10, 14]) and its increasing availability demand the development of efficient protocols to familiarize operators with remote environments. VR is a promising candidate for this task because of its increasing affordability and capabilities to communicate spatial information. Meanwhile there are efforts to cost-effectively produce virtual models of buildings [15–17]. All the necessary components for effective use of VR to prepare teleoperation in remote environments are available

- *Karl Eisenträger is with Humboldt-University Berlin. E-mail: karl.eisentraeger@protonmail.de.*
- *Judith Haubner and Georg Jahn are with Chemnitz University of Technology. E-Mail: {judith.haubner | georg.jahn} @psychologie.tu-chemnitz.de.*
- *Jennifer Brade, Sven Winkler and Philipp Klimant are with Chemnitz University of Technology. E-mail: {jennifer.brade | sven.winkler | philipp.klimant} @mb.tu-chemnitz.de.*
- *Wolfgang Einhäuser and Alexandra Bendixen are with Chemnitz University of Technology, E-Mail: {wolfgang.einhaeuser-treyer | alexandra.bendixen} @physik.tu-chemnitz.de.*

Manuscript received 14 October 2022; revised 13 January 2023; accepted 30 January 2023. Date of publication 22 February 2023; date of current version 29 March 2023. Digital Object Identifier no. 10.1109/TVCG.2023.3247052

already and are likely to experience growth in use. Hence, we used a teleoperation task as our application scenario for testing how different types of VR-based learning compare to map-based learning of an unfamiliar environment.

1.3 Test setting

The unknown and remote environment we used for testing the feasibility of VR as a preparation tool, was a floor in a university campus building that had been faithfully reconstructed in the context of earlier work on gaze allocation in VR [18]. With its variety of corridors and more open spaces, we consider it representative of a large class of buildings. We employed two different VR methods, one with the user's avatar taking the perspective of the robot that was later going to be teleoperated, the other with a giant-sized avatar. As reference, we used the standard method for communicating the layout of a building floor – the floor plan as a 2D map. After training with exactly one of these methods, all participants performed the same orientation and navigation tasks by operating an actual telerobot remotely through the environment.

2 Related Work

Performance in many spatial tasks is impaired in synthetic, mediated environments such as VR compared to the real world (e.g.,[19]). Prior work has identified orientation and navigation as key challenges when operating a telerobot [12, 13]. VR has been suggested as a direct control mechanism to operate telerobots [20–25]. Our present study in turn focuses on environment learning in VR for subsequent teleoperation in the real environment. In everyday life, humans rarely experience discontinuity when it comes to where they are in space. When initiating the teleoperation, in contrast, the operator must gauge the approximate location of the robot in the remote environment to be able to effectively navigate and operate. We will refer to this demand as initial orientation. Navigational orientation and initial orientation are both needed for successful use of a telerobot. In one study it was observed that switching perspectives from learning to testing resulted in costs and that recognition performance was enhanced if the egocentric orientation was the same in test and learning, but only if the perspectives matched [26]. This corresponds to the more general finding in studies examining environment learning that performance is better when the learning perspective matches the test perspective [26–29].

2.1 Environment learning

Environment learning is a process in which humans encode spatial information regarding relationships of locations within an environment [5]. This learning can take place through a survey perspective (top-down view, learning from maps) or from within the environment or with both combined [30]. The resulting mental representations differ between the learning methods [31, 32]. Learning with maps provides an advantage for estimating global spatial relations and straight-line distances between locations, whereas learning by navigating inside the environment benefits assessing self-to-landmark spatial relations [30]. Thus, the utility of survey perspectives depends on the demands of the chosen navigation tasks. For instance, one study investigated the learning of an unfamiliar campus building by a map versus navigating within the building and found that map learning benefitted survey knowledge, whereas navigation learning led to better performance on navigation tasks assessing route knowledge [33]. Another study comparing four groups learning a virtual environment (2D) through videos of a simulated journey - in a survey perspective that resembled an aerial perspective, a first-person perspective, both perspectives combined, or no video at all - concluded that the survey perspective supported navigation on unexpected detours and far-space navigation better than the first-person perspective [34]. The first-person perspective in this case supported only a restricted range of local navigation. The "survey" conditions in these two studies differed considerably: The map used to show the campus building in [31] was abstract, monochrome and only featured the raw layout of the building floor and the names of the rooms therein. The survey view employed in the video of the simulated journey in [34] on the other hand, was not abstract ("just" a shift in perspective), multicolored, featured the names of locations as well, and was a video instead of a static image [34]. To directly compare a map and an elevated perspective for survey, we here studied these different methods of environment learning with tasks that were identical for all learning conditions.

VR can lead to enhanced performance and affords new ways in which machines can be teleoperated [20–25]. One study allowed the user to control an unmanned aerial vehicle and an unmanned ground vehicle at the same time as control of the vehicles was semiautonomous [20]. The two vehicles reconstructed their surroundings as VR and streamed a VR model to the user. The user could take on different perspectives in the streamed VR and shrink or enlarge their avatar similar to multiscale virtual locomotion [35]. A related paper proposed VR prototyping-based path planning for unmanned aerial vehicles for the inspection of building exteriors [36]. In the process a virtual model was uploaded into a VR environment, and via an unmanned aerial vehicle simulator plugin (allowing for a first-person view and a third-person view onto the drone), an expert would repeatedly maneuver a drone until satisfaction with one of the flight paths was reached. In contrast to studies using VR for active operation, semi-autonomous operation, or prototyping-based path planning, our present study focuses on environment learning in VR for subsequent active teleoperation in the real environment.

2.2 Giant Virtual Reality (GVR)

Argelaguet and Maignant used the acronym "GiAnt" (Giant/Ant) to describe a multi-scale navigation technique to explore virtual environments [37]. The technique automatically adjusts the scale factor of the virtual environment according to the perceived navigation speed of the user (giant perspective for fast navigation). Around the same time the team of Google Earth VR presented their take on perspective and navigational mode, which had similar mechanics to "GiAnt" and scales the users' avatar size while flying according to their eye level in the scene [6]. Drastically changing the avatar size to convey the feeling of standing on the ground was presumed to reduce cybersickness and fear of heights experienced by users. This would allow the user to change their perspective freely in terms of height while the automatic adjustment of avatar size would evoke the sense of looking at a miniature world rather than looking from far away onto the world. This is achieved by adjusting the vergence angle: While increasing the avatar size, the modeled eye distance is increased as well, providing a larger vergence angle [6]. A study investigating different methods to enable users to rapidly explore virtual environments included "Ground-Level Scaling", "Eye-Level Scaling", and "Seven-League Boots" [38]. "Ground-Level Scaling" describes the enlargement of the player avatar from ground level. With "Eye-Level Scaling" the avatar's size is also increased but the level of the eye height is maintained at the height of a normal-sized avatar, thereby the eye-level is maintained but the perspective is changed through the increased vergence angle, and the movement speed is increased linear to the increase in avatar size. Both enlargement methods result in a speed gain linear to the gain in size as well as in a linear increase in modeled eye distance and thereby in the perception of the virtual environment as a miniature. The "Seven-League Boots" simply amplified the user's movement. For the exploration task, participants preferred "Ground-Level Scaling" over the other two methods; this method also resulted in higher scores of embodiment and allowed the users to maintain positional accuracy and control even at high speed gains [38]. Another study explored a navigation technique that switched between a giant-sized and a normal-sized avatar and compared it to teleportation in their ability to enable fast traveling and seamless orientation in a videogame set-up [39]. The study found an increase in spatial orientation while avoiding cybersickness and maintaining presence, enjoyment, and competence.

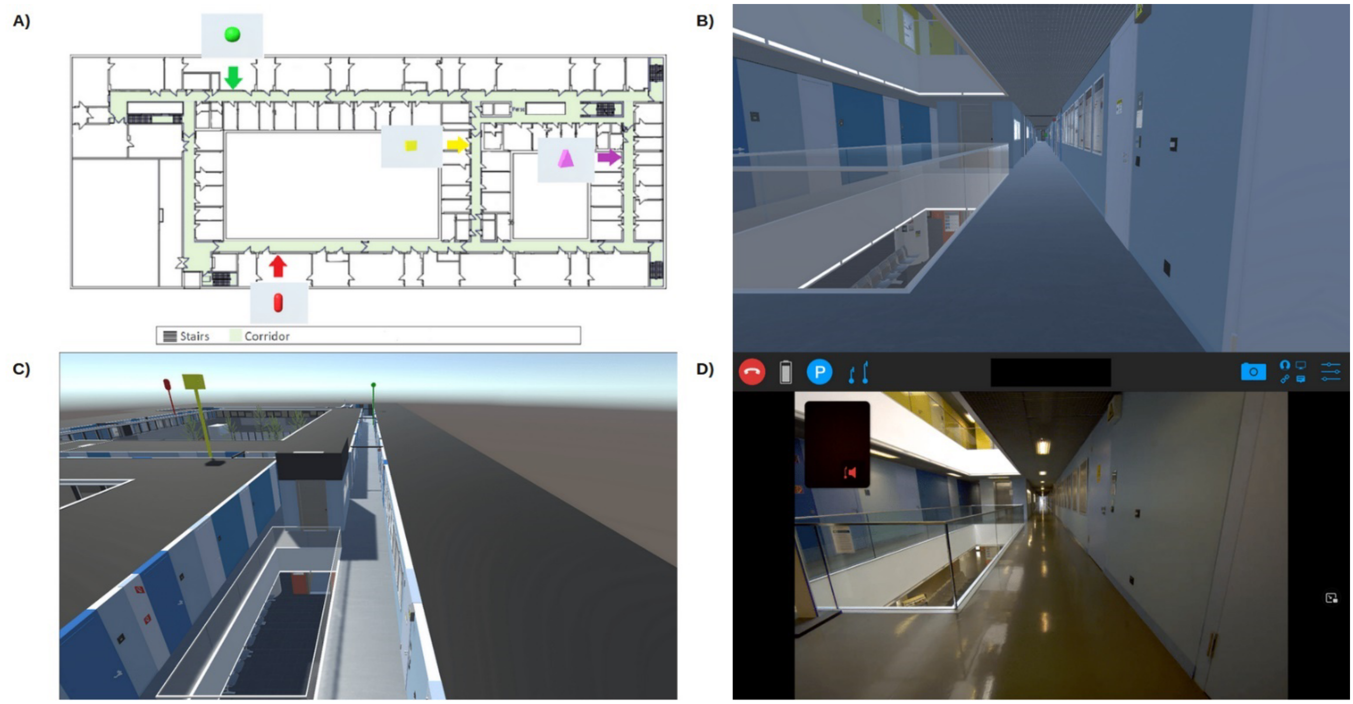

Fig. 1. (A) Building plan as used by the BP group (for the actual experiment, the legend was provided in German). (B) VR as seen by the NVR group. (C) VR as seen by the GVR group; note the checkpoint markers put on poles. (D) Double 3 interface as seen by all participants during the initial orientation and navigation tasks. (B)-(D) show views from approximately the same map location, which corresponds to a point on the upper right in the map of panel (A).

2.3 Related VR perspectives

The use of different VR perspectives to aid orientation in virtual spaces has been researched in a variety of contexts. One study primarily concerned with occlusion management in indoor spaces compared conventional view, map view, X-ray visualization and multiperspective visualization in regard to spatial awareness, depth perception, and performance metrics [40]. The multiperspective visualization is a continuous, non-redundant integration of samples from multiple viewpoints into one image that reduces occlusion. In two tasks that required the participants to find items/persons in two environments, the multiperspective visualization outperformed the other perspectives. Another study employed a perspective in virtual reality called “UrbanRama” -a cylindrical projection that integrates a navigational and a survey perspective in one view- and regarded “UrbanRama” to be comparable in effectiveness to aid navigation and orientation in virtual cities to bird’s-eye view and to a mini-map [41]. Presumably due to the relatively small number of participants (N = 10 per group) only trends could be observed; these seemed to show an advantage of “UrbanRama” over the other two methods for seasoned VR users. In comparison to these perspectives, we regard GVR as a simpler approach that relies only on a change of scale and therefore might be more readily usable, understood, and implemented.

2.4 Cybersickness

Cybersickness describes a subtype of motion sickness that can be triggered when humans experience VR [42–44]. There are many factors influencing the likelihood and severity of cybersickness including the hardware used, content of the virtual scene, duration of the time spent in VR, and individual factors [43, 45]. Central symptoms of cybersickness are disorientation, nausea, and eye fatigue [42]. Because cybersickness is a very unpleasant experience, it may reduce the time humans spend in VR and thereby reduce the usefulness of VR. This fact, in line with general ethical considerations, commands reducing cybersickness as much as possible.

3 Our Approach and Hypotheses

We investigated how different methods of training influenced subsequent initial orientation and navigation in an unfamiliar environment. To this end, we pseudo-randomly assigned participants to three training groups. One group used a building plan (BP, Fig. 1A), the other two used VR. One of the VR groups took the perspective of a normal-sized avatar (NVR, Fig. 1B), the second VR group took the perspective of a giant-sized avatar (GVR, Fig. 1C). After this learning phase, all participants conducted the same spatial tasks with a teleoperated robot: initial orientation and navigation. We measured the time that participants decided to use for training, the correctness in the initial orientation task, and the time to complete the navigation task. In addition, we assessed the sense of direction for the participants of all groups and self-reported cybersickness for the two VR groups.

For the training phase, we expected that the time taken for learning would be greater in the NVR condition than in the GVR condition as the GVR condition should allow for faster exploration and greater oversight of the VR environment [38, 46]. We expected the learning to be the shortest for the BP group as no navigation is necessary and the spatial information is abstract and can readily be accessed. We furthermore expected that there would be a lower cybersickness rating for the GVR group compared to the NVR group as a consequence of the expected difference in exposure times, while no extra burden is expected to arise due to the different perspective in GVR [6, 38, 42, 46].

For the initial orientation task, we expected the NVR and GVR groups to show a better performance than the BP group because navigational learning is expected to benefit self-to-landmark spatial relationships [30]. Self-to-landmark spatial relationships are one important element of what we coined initial orientation (yet they do not reflect the aspect of discontinuity). We expected a benefit of NVR over GVR in terms of performance as its perspective was identical to the test perspective, which should support recognition [26–29].

For the navigation task, we expected that NVR and GVR would outperform BP because a similar study indicated that navigation would support route knowledge better than map studying [33]. We also expected an advantage of GVR over NVR as it employs a navigational and a survey perspective at the same time, which was found to benefit orientation in a range of earlier studies [34, 38, 46].

4 Methods and Materials

4.1 Participants and setting

Sixty-six adults participated in the study, 22 in each of the three experimental conditions. All participants gave written informed consent. They had not been to the physics building of Chemnitz University of Technology nor did they possess any knowledge about its structure. Participation was compensated with money or study credits. Normal or corrected-to-normal vision was a further inclusion criterion for participation. Although this criterion was communicated clearly during recruitment, it became evident based on their responses in the demographic questionnaire that three participants (2 of the NVR, 1 of the BP group) did not meet this requirement. These were excluded from all analyses; the demographics of the included participants are shown in Tab. 1. The experiment was conducted simultaneously in two places – Chemnitz University of Technology and Humboldt University Berlin, which are located more than 100 miles apart. Participants were in a lab on the university campus of Humboldt University Berlin where they were instructed and monitored throughout the experiment by the experimenter. At the same time another experimenter set up the robot for the two tasks at Chemnitz University of Technology. This created an actual teleoperation situation and ensured that the participants were indeed unfamiliar with the environment the telerobot operated in.

Table 1. Demographic Information (Participants Included in the Analysis)

	Gender	Age	VR experience (own VR equipment)
BP ($N = 21$)	13 men, 8 women	19-59 ($M = 31.52$, $SD = 10.73$)	8 (1)
NVR ($N = 20$)	4 men, 16 women	20-57 ($M = 31.35$, $SD = 9.68$)	8 (1)
GVR ($N = 22$)	7 men, 16 women, 1 non-binary person	19-36 ($M = 28.7$, $SD = 4.7$)	6 (1)

4.2 Learning scenarios

There were three learning scenarios, one for each group - BP, NVR, and GVR. The BP represents the standard approach to environment learning, while NVR and GVR represent virtual alternatives to be compared to this standard.

4.2.1 Hard– and software setup

The computer used for all tasks was a ThinkPad Workstation (Intel i7 processor, Nvidia Quadro RTX3000). The main screen used to display the questionnaires and the robot operational interface was a BENQ BL2780 with 27.2 inches diagonal. The IBM laptop screen was used in the orientation task to display the response options. An Oculus Quest 1 Head-Mounted Display (HMD) was used for the VR experience. For interaction the Oculus Quest 1 controllers were used.

The VR scenario was handled in Unity 3D (version 2021.1.20f1) [47]. It consisted of a model of a building [18]. The model (Fig. 1B, and Fig. 1C) resembled in detail the architecture, furnishing (trashcans, printer, pallet truck, posters) and windows of the real building (Fig. 1D). In the virtual scene, four locations were highlighted with differently colored geometric 3D-forms. These geometric forms were referred to as checkpoints and served two purposes: First to mark a certain location and second to identify it (e.g., "green sphere").

The "XR Interaction Toolkit" from Unity was used to provide locomotion. Pitch, yaw, and roll movements were enabled for the HMD-tracking (3 degrees of freedom). The locomotion was set to continuous movement and could be triggered with the controllers. The joystick of the left controller could be used to move forwards, backwards, left, and right. The joystick of the right controller could be used for leftward and rightward turns.

4.2.2 Normal virtual reality (NVR)

The NVR learning scenario featured the environment to be learned. The camera height (119cm) and locomotion speed (3.6 km/h) were defined according to the properties of the teleoperated robot. The controllers as well as the ray interactive lines were not visible in VR (Fig. 1B). Besides moving in the environment no interactions were possible. There was no sound added to the experience. Doors could not be opened; elevators and staircases could not be accessed - effectively limiting the free movement to the corridors of the floor.

4.2.3 Giant virtual reality (GVR)

The GVR environment was based on the same model as the NVR environment. Everything but the user avatar was downsized by a factor of 5. Thereby the user effectively became a giant. The movement speed remained the same as for NVR, such that the user effectively moved five times faster relative to the building. The accessibility of the model remained the same. Some parts of the environment had to be changed to accommodate the new perspective. The ceiling of the first floor and everything above it was removed from the scene. These removed parts would have obscured the vision of the user. Parts of the model that were not relevant for the experiment and could be seen in GVR vs. NVR (e.g., adjunct rooms that were irrelevant for the experiment) were covered with black planes preventing distraction. To maximize the visibility of the checkpoints, they were put on poles with the same coloration and raised to eye-height of the user (Fig. 1C).

4.2.4 Building floor plan (BP)

We used the pre-existing building plan which was accompanied by a legend (Fig. 1A). The corridors to be learned were highlighted and the room numbers were removed – as they were from the door signs in the VR conditions. The checkpoints were highlighted with colored arrows pointing towards their respective locations. The legend entailed the colored geometric forms. The building plan was presented on a screen at fixed size and resolution.

4.3 Teleoperated robot

The robot used for the orientation and navigation task was the Double 3 [48] (Fig. 2). The Double 3 is a telepresence robot with a self-balancing wheel, stereo vision depth sensors, ultrasonic range finders, wheel encoders, and an inertial measurement unit. The body frame is 25cm x 33cm x 119-150cm (W x D x H). The height was set to the lowest possible setting (119cm) as the movement speed is highest at that setting. "Mixed Reality Settings" and "Obstacle Avoidance" were turned off. The former was turned off to avoid distortion of the camera stream. The latter was turned off as it incorrectly recognized light reflections on the floor as obstacles - thereby rendering effective navigation impossible when turned on. The screen brightness setting was on standard. The interface screen of the Double 3 shows the live stream of the cameras positioned on its "head". Movement was controlled through either the "W", "A", "S", and "D" keys or the arrow-keys on the keyboard. This was the main input control for operating the robot. The other function used for the navigation task was the screenshot function, which was triggered by clicking on an icon depicting a photo camera (Fig. 1D). No video or audio was streamed to the robot. Audio from the robot to the participant was streamed to be able to react to an emergency. Between trials the Double 3 was moved to the starting position of the next task by the second experimenter. During the transfer its camera was covered, and the microphone turned off.

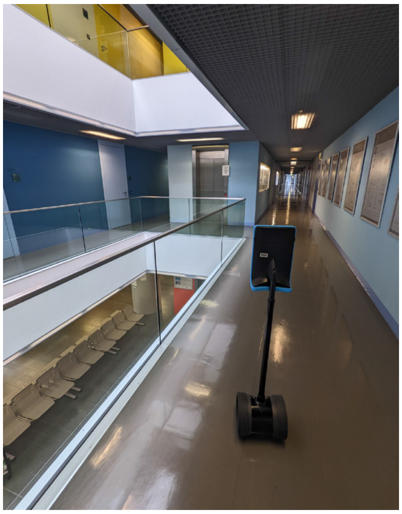

Fig. 2. Double 3 in its lowest height (119cm) in one of the corridors used for the experiment seen from the back.

4.4 Procedure and measures

An overview of the procedure is depicted in Fig. 3.

Initially, the participants gave written informed consent, signed the privacy agreement, filled out a questionnaire on demographic information (age, gender, eyesight, prior VR experience), and answered the Santa Barbara Sense of Direction Scale (SBSOD) [49].

In the following segment, the participants learned to operate the telerobot in a secluded room such that the environment to be learned remained invisible. The exercise required them to navigate along a short route marked on the floor, take a picture of an A4 print of a checkpoint with the integrated screen-shot function, take a turn and steer back to the starting position.

All groups were informed about the two tasks they would have to perform with the telerobot. For the VR groups, a brief familiarization session for the movement control in VR (in an unrelated VR environment) and a first Simulator Sickness Questionnaire (SSQ) [50] assessment was added before the learning. Subsequently, all participants were instructed to learn the environment with their respective methods (BP, NVR, GVR) until they felt confident and prepared. The goal of the learning was to be able to recognize the locations without the checkpoint markers, and to be able to navigate between them. The learning itself was free and had no time constraint - but the utilized time was captured. The VR groups were again assessed with the SSQ after the learning for a post measurement. Whenever a participant experienced discomfort after the VR experience, we offered water, chewing gum, and dextrose as well as a break to counteract the effects of potential cybersickness. The break did not count as part of the learning time. In these occurrences, the experimenter also explicitly reminded the participant that they could abort the experiment at any time without any repercussions.

Thereafter, all groups were asked to perform the eight trials of the initial-orientation task. In preparation of the trials, the second experimenter at Chemnitz University of Technology moved the telerobot to the pre-planned locations, which were evenly distributed in terms of closeness to the checkpoints and the telerobot's initial alignment relative to the closest checkpoint. The participant had to answer which of the checkpoints was closest to the initial position of the telerobot in each trial[1]. Responses had to be indicated on a second screen, with the mouse cursor positioned in the middle of the response options at the beginning of each trial. A beeping sound indicated to the participant the start of the trial and revealed the telerobot video stream. To support orientation, the participants could move the telerobot freely. Each trial had a time limit of 60 seconds, with a countdown depicted on the second screen and a second beeping sound 15 seconds before the end of each trial. If the question was not answered at the expiration of the time (60 seconds in), the trial was stopped. The time to respond and the response itself were captured; not responding within the time limit was counted as an incorrect response.

The second task was a navigation task, which encompassed four trials. For this task the checkpoints were marked with colored A4 prints of the geometric forms associated with them. The experimenter at Chemnitz University of Technology again moved the telerobot to the starting positions for each trial. The starting positions were right in front of the checkpoints with the camera pointing towards the checkpoint. The participants were then verbally instructed to navigate the telerobot to another preselected checkpoint on the shortest route. Upon revelation of the telerobot video stream, the participants were asked to take a screenshot and upon arriving at the target checkpoint to take another one. Participants were instructed to take the screenshots with the telerobot being right in front of the A4 prints. The timestamps of the screenshots were used to calculate the navigation time.

At the end, the participants were asked to answer questions related to their experience, give feedback, describe strategies they used to solve their tasks, and indicate whether they had participated in similar experiments before.

All procedures were determined by the applicable body (Ethics committee of the Faculty of Behavioral and Social Sciences of Chemnitz University of Technology) not to require in-depth ethics evaluation (V-332-15-GJ-Telepresence-13052019).

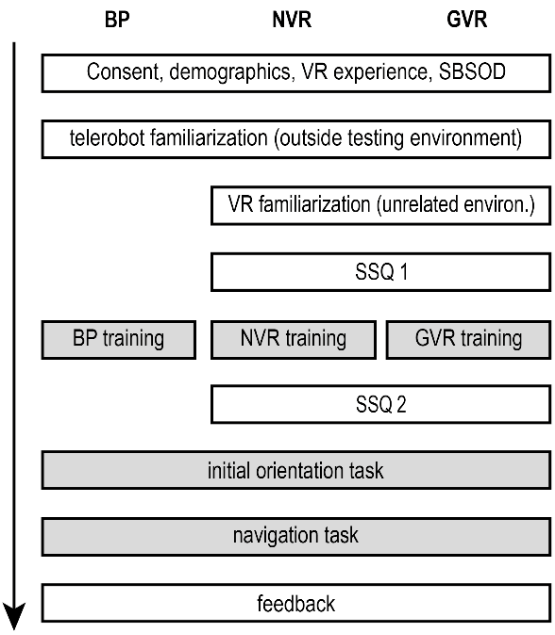

Fig. 3. The procedure of the experiment.

[1] Although the task uses distance estimation we use the term "orientation task" to express that users have to orient themselves in (i.e., have to get into) the novel situation; it is not meant in the sense of acquiring a particular orientation (i.e., heading).

5 Results

Although the assignment to groups was random, we first checked whether there were any differences evident in their sense of direction according to the SBSOD scale using its original formula [49]. We found average values of 4.49 (SD =1.01) for the BP group, 3.89 (SD = 1.18) for the NVR group and 4.21 (SD = 0.75) for the GVR group. As we have no evidence for a violation of normality of the distribution in any of the groups (Shapiro-Wilk test all $p > .134$) nor for a violation of variance homogeneity ($F(2,61) = 2.63$, $p = .134$, Levene's test), we conducted a one-way analysis of variance (ANOVA) to assess group differences. We found no evidence that the SBSOD differed between the groups ($F(2,61) = 1.90$; $p = .157$).

We assessed whether the subjective feeling of cybersickness changed differently for the two VR methods during the VR training. To this end, we computed the difference in SSQ between after and before the training (Fig. 4), where the SSQ computation followed the original definition [50]. For the distribution of differences, the assumption of normality was violated for the NVR ($W = 0.89$; $p = .037$, Shapiro-Wilk) and for the GVR group ($W = 0.84$; $p = .012$). Therefore, we carried out a non-parametric Mann-Whitney test to test whether there was evidence for a difference between the groups. We found no indication that the group affected the difference in SSQ total score between the pre- and post- assessment ($U = 256$; $p = .369$).

Learning time evidently was not normally distributed for the BP group (Fig. 5). Therefore, we used the Kruskal-Wallis non-parametric test to test whether learning time depended on the group. We indeed found an effect of group ($H(2) = 33.83$, $p < .001$, $\eta 2 = .53$, $d = 2.13$). A post-hoc Dunn's test indicated that the NVR group differed significantly from the GVR ($p < .001$) and the BP ($p < .001$) groups at a Bonferroni-corrected alpha level (corrected alpha = .05/3 = .017). We did not find a difference between BP and GVR ($p = .519$).

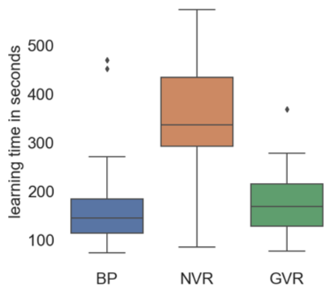

Fig. 5. Boxplots of learning time in seconds for BP, NVR, and GVR.

For the initial orientation task, we considered the number of correct responses over the eight trials, where timeouts, which accounted for 6.35% of responses (11.31% BP; 5.63% NVR; 2.27% GVR), were counted as incorrect. We summed the correct responses per participant. Since – due to the discrete nature of the measure – a normal distribution cannot be expected, we used the Kruskal-Wallis non-parametric test to assess the effect of group. We found that the amount of correctly identified checkpoints in the orientation task differed significantly between groups ($H(2) = 24.82$, $p < .001$, $\eta^2 = .38$, $d = 1.567$; Fig. 6). Post-hoc Dunn's tests indicated that the BP group differed significantly from the GVR group ($p < .001$) as well as from the NVR group ($p = .007$). The difference between the NVR and the GVR groups ($p = .028$), albeit significant at a .05 alpha level, does not reach significance when a Bonferroni correction for the three independent comparisons is applied and the alpha-level therefore adjusted to .017 (.05/3).

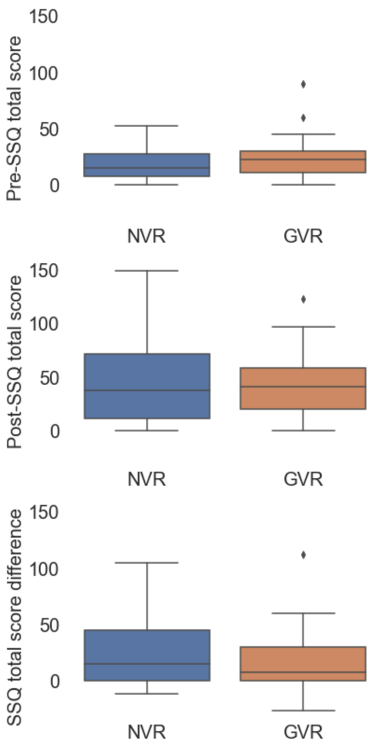

Fig. 4. Boxplots for SSQ pre-, post- and difference score. Lower and upper bound of box indicate 25% and 75% quartile, respectively, horizontal line the median, whiskers extend only up to farthest data-point present in the 1.5 interquartile range, values outside the whiskers are represented individually.

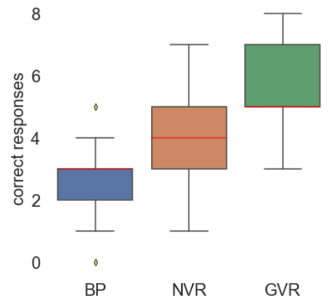

Fig. 6. Boxplots of the number of correct responses in the initial orientation task for BP, NVR, and GVR. Here the median lines are colored red for visibility as they – due to the discrete nature of the variable – coincide with the 75% percentile (BP) and the 25% percentile (GVR), respectively.

For the navigation task, we accumulated the time per trial to a total time (Fig. 7). As for the distribution of differences, the assumption of normality was violated for the NVR ($W = 0.89$; $p = .032$, Shapiro-Wilk) and for the GVR group ($W = 0.85$; $p = .004$). We applied a Kruskal-Wallis test to assess whether the total navigation time depended on the group. We found an effect of group on total time ($H(2) = 13.59$, $p < .001$, $\eta^2 = .19$, $d = 0.98$). Post-hoc Dunn's tests indicated that the GVR group differed significantly from the BP group ($p < .006$) and from the NVR group ($p < .001$) at a Bonferroni-corrected alpha level of .017. We did not find a difference between BP and NVR ($p = .460$).

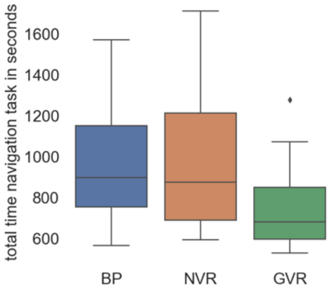

Fig. 7. Boxplots of total time in seconds used in the navigation task for BP, NVR, and GVR.

6 Discussion

Our study compared two different kinds of perspectives in VR and the use of a building plan as preparation tools for orientation and navigation of a telerobot in an unfamiliar environment.

6.1 Learning time

Participants were instructed to decide themselves when to end the preparation (environment learning) phase as they felt ready for completing the teleoperation tasks. Despite substantial inter-individual variability, we found a clear pattern in that participants spent more time in NVR than in either GVR or BP, which were about equal on average. This may indicate that GVR and BP constitute more time-efficient learning tools than NVR for an environment of our scale. Our finding is in accordance with other research concerned with VR interactions that take similar forms as our GVR [6, 38, 46]. Two properties of GVR likely contribute to this advantage over NVR: first, a larger part of the environment can be inspected at once; second, the user can explore the environment faster due to the increased speed relative to the building size (fast vs. slow translation speed in VR did not affect orientation in [51].

6.2 Cybersickness

When systematically assessing cybersickness scores, we did not find evidence for an effect of group on the difference between SSQ scores before and after the VR experience. This is contrary to our initial expectation that the longer exposure time of the NVR compared to the GVR might result in higher SSQ scores. Most likely, any subtle increase in cybersickness saturates quickly after the onset of the VR experience, such that the exposure time plays only a minor role in our specific setting. It should be noted that in absolute terms the SSQ scores seem rather high at first glance [52]. However, this already applies to the pre-test scores that were recorded after a very brief familiarization phase. The most probable explanation for these high SSQ scores before any substantial VR exposure is that the experiments were conducted late in the evening (7pm+) to minimize the occupation of the building the robot was operated in. The values are well within a range reported as baseline values in a recent reassessment of the SSQ [53] and the assumption of a near zero-baseline - albeit probably reasonable in the aviation context the SSQ had originally been developed for - might be inappropriate for assessing cybersickness with head-mounted displays in the average population [50, 52]. This is in line with the observation by the experimenter who surveilled the participants throughout the experiment and in most cases saw no indication of cybersickness. Even in the few cases where some discomfort was reported by the participants, they were eager to continue with the main experiment after a short break and to see some further VR demos after the end of the experiment. In sum, it is exceedingly unlikely that the exposure to VR in the present context causes substantial amounts of cybersickness – at least for most individuals. Nonetheless, keeping the VR training protocol for navigation tasks as efficient as possible to minimize the needed VR exposure time for training will always be preferable. Comparing different VR methods, as we do herein, contributes to this overarching objective.

6.3 Initial orientation

For the initial orientation task, GVR and NVR both outperformed BP. This part of the result pattern was not surprising and in line with research indicating that similarity between the study perspective and the test perspective leads to better performance [26, 28]. Following the same line of argument, another part of the results of the initial orientation task was surprising, as GVR showed a trend to outperform NVR. NVR more closely resembled the testing perspective and at first glance we therefore expected it to be more effective at aiding initial orientation. This would also have been in line with previous research that found self-to-landmark orientation to be better when learning in the environment itself compared to learning with a map [30]. However, in our experiment, the input from the robot was already substantially different from both VR conditions (and from being in the environment itself) – for example, the video stream was two-dimensional and presented on a screen compared to the 3D VR presented in an HMD. Maybe this difference in display already diminished the association that would benefit NVR to such an extent that other factors favoring GVR dominated the result pattern. In GVR, participants perceived the checkpoints and their surroundings at once due to their elevated perspective. This may have helped them to create a better integrated and more exhaustive mental representation of the checkpoints, their relative position, and their closest surroundings, especially as other clearly discernible landmarks were scarce in the specific environment.

6.4 Navigation task

The results of the navigation task were partly in line with our expectations. GVR outperformed both NVR and BP, but NVR did not perform better than BP. Because GVR combines a navigational and a survey perspective [34] and is associated with high orientation gains [46], we expected it to perform the best - which it did. As the GVR perspective provided the relative locations from the checkpoints to each other in one view from certain angles of the virtual environment, it may have been easier for the participants in that group to adequately plan routes and identify relevant landmarks helping their performance later on. However, NVR performed roughly at the same level as BP. This was unexpected because previous literature would suggest that the navigational perspective should have supported subsequent navigation better than the sole survey perspective provided by BP [30, 33]. For the NVR group it must be mentioned that the movement speed (3.6 km/h) was slow because it was set according to the movement speed of the robot. Some of the participants indicated that the movement speed was too slow for them and asked for faster movement. The slow movement speed might have prevented participants from thoroughly exploring the virtual environment and thereby explain the performance of the NVR group, including its comparably larger inter-individual variability.

7 Limitations

A range of limiting factors can be identified in our study. One was the relative inexperience of participants with respect to VR: Most participants used VR for the first time (66.67%), only two of them owned VR equipment. The teleoperation of the robot was also new to almost all the participants and the steering abilities of the participants differed. This was unexpected to us because in test trials we did not observe the variance we later observed in the experiment. Another related issue pertains to the use of two screens simultaneously as was done in the initial orientation task. This may have drawn on multi-tasking skills unrelated to the target variables of this study. Inter-individual variation in such skills, if not distributed evenly among the groups, might have obscured some effects of interest. A reoccurring complaint from the participants concerned the low movement speed of the robot. The low movement speed did level the influence of steering ability but at the same time might have obscured additional differences in route finding ability because the low movement speed led to ample time available for route contemplation. Moreover, the environment we used was a rather monotonous one with few landmarks and its layout was rather regular. It is quite possible that an environment featuring more landmarks and a more diverse structure might benefit from learning in VR more as it would play more to its advantages. A final, more curious limitation can be seen in the fact that some of the participants thought that the teleoperation of the robot was a simulation of some kind. This happened although we informed the participants about the procedure of the study at several points in time (study announcement, participant information, repeatedly during the experiment) and supplemented information material with pictures of the robot.

8 Conclusion and Future Work

In this paper, we investigated the use of VR as a preparation for teleoperation in unfamiliar environments. We compared a standard building plan to two different perspectives in VR. To test the viability of the different approaches, we measured the performance of participants in two different tasks drawing on (1) orientational and (2) navigational capabilities.

The results of our experiment confirmed our notion that VR is a useful and effective tool to prepare teleoperation in unfamiliar environments. Specifically, we could extend findings of other studies relating to the benefits of GVR for orientation and navigation into a new context [6, 35, 37, 38, 46]. We also observed that VR supports initial orientation with a teleoperated robot, which might be useful in cases where no humans are present in the remote environment, quickness to act is important, and when the position of the robot in the environment is unknown.

Another aspect relevant for teleoperation of robots is the spatial information that can be communicated in VR about the absolute dimensions of the robot that will be teleoperated and its dimension relative to the environment. In our experiment, this was not relevant since the robot we employed is slim and small and there were no challenging passages in the environment. In general, this may not be the case and knowledge about difficult passages might be deduced in VR and may be prepared for in VR.

For future research, it would be advisable to choose a robot for teleoperation that operates at a higher speed than the one we used to better tease out differences in spatial performance and prevent impatience in the participants. Additionally, it would be preferable to have the same interaction in VR and with the robot in terms of steering.

Moreover, it would be interesting to measure the user-experience of the participants regarding different perspectives in VR. To enable the participants to switch between perspectives (or adjust them) would allow for observation of preferences and identification of the situations which make participants prefer a certain perspective. It would also be interesting to allow for different speeds to discern effects of perspective from effects of speed. In a similar vein, it would be appealing to discern at what scale of the environment the positive effects of GVR set in. To test GVR across different use cases and environment sizes against other perspectives, modes and interactions such as “world-in-miniature”, “UrbanRama” or multiperspective visualization [40, 41, 54–56] might help users and developers to make informed choices using and developing VR.

Acknowledgments

We thank Jenny Rettstatt for supporting data acquisition. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 416228727 – SFB 1410.

References

[1] D. Harley, “Palmer Luckey and the rise of contemporary virtual reality,” Convergence, vol. 26, no. 5–6, pp. 1144–1158, Dec. 2020, doi: 10.1177/1354856519860237

[2] Z. Duer, L. Piilonen, and G. Glasson, “Belle2VR: A Virtual-Reality Visualization of Subatomic Particle Physics in the Belle II Experiment,” IEEE Computer Graphics and Applications, vol. 38, no. 3, pp. 33–43, May 2018, doi: 10.1109/MCG.2018.032421652

[3] M. D. Lopez, “How can I get there? Disorientation and motion sickness effects depending on the virtual reality navigation technique,” Master’s Thesis, Aalborg University, Denmark, 2020. [Online]. Available: https://projekter.aau.dk/projekter/en/studentthesis/how-can-i-get-there-disorientation-and-motion-sickness-effects-depending-on-the-virtual-reality-navigation-technique(970345a5-123d-4e3d-a3da-911900a57a53).html

[4] E. Olshannikova, A. Ometov, Y. Koucheryavy, and T. Olsson, “Visualizing Big Data with augmented and virtual reality: challenges and research agenda,” Journal of Big Data, vol. 2, no. 1, p. 22, Oct. 2015, doi: 10.1186/s40537-015-0031-2

[5] R. Dai, A. K. Thomas, and H. A. Taylor, “When to look at maps in navigation: metacognitive control in environment learning,” Cogn Res Princ Implic, vol. 3, p. 36, Sep. 2018, doi: 10.1186/s41235-018-0130-7.

[6] D. P. Käser et al., “The making of Google earth VR,” in ACM SIGGRAPH 2017 Talks, New York, NY, USA, Jul. 2017, pp. 1–2. doi: 10.1145/3084363.3085094

[7] R. G. Boboc, H. Moga, and D. Talaba, “A review of current applications in teleoperation of mobile robots,” Bulletin of the Transilvania University of Brasov. Engineering Sciences. Series I, vol. 5, no. 2, p. 9, 2012.

[8] L. Basañez and R. Suárez, “Teleoperation,” in Springer Handbook of Automation, Springer, 2009, pp. 449–468. [Online]. Available: 10.1007/978-3-540-78831-7_27

[9] L. Cheng and M. Tavakoli, “COVID-19 Pandemic Spurs Medical Telerobotic Systems: A Survey of Applications Requiring Physiological Organ Motion Compensation,” Frontiers in Robotics and AI, vol. 7, 2020, doi: 10.3389/frobt.2020.594673

[10] T. B. Sheridan, “Telerobotics,” Automatica, vol. 25, no. 4, pp. 487–507, Jul. 1989, doi: 10.1016/0005-1098(89)90093-9

[11] A. V. Gomez, C. Saracini, F. Botta, and G. Randelli, “Spatial processes in mobile robot teleoperation,” Cognitive Processing, vol. 10, no. S2, pp. 338–341, 2009, doi: doi: 10.1007/s10339-009-0299-2

[12] S. Opiyo, J. Zhou, E. Mwangi, W. Kai, and I. Sunusi, “A Review on Teleoperation of Mobile Ground Robots: Architecture and Situation Awareness,” Int. J. Control Autom. Syst., vol. 19, no. 3, pp. 1384–1407, Mar. 2021, doi: 10.1007/s12555-019-0999-z

[13] J. Chen, E. Haas, and M. Barnes, “Human Performance Issues and User Interface Design for Teleoperated Robots,” Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 37, pp. 1231–1245, Dec. 2007, doi: 10.1109/TSMCC.2007.905819

[14] L. Du, J. Yuan, S. Bao, R. Guan, and W. Wan, “Robotic replacement for disc cutters in tunnel boring machines,” Automation in Construction, vol. 140, p. 104369, Aug. 2022, doi: 10.1016/j.autcon.2022.104369

[15] D. Santos, M. Dion’isio, N. Rodrigues, A. Pereira, and I.-I. Leiria, “Efficient Creation of 3D Models from Buildings’ Floor Plans,” International Journal of Interactive Worlds, Nov. 2011, doi: 10.5171/2011.897069

[16] A. Ammar et al., “Digital Twins in the Construction Industry: A Perspective of Practitioners and Building Authority,” Frontiers in Built Environment, vol. 8, 2022, doi: 10.3389/fbuil.2022.834671

[17] E. Lewandowicz, F. Tarsha Kurdi, and Z. Gharineiat, “3D LoD2 and LoD3 Modeling of Buildings with Ornamental Towers and Turrets Based on LiDAR Data,” Remote Sensing, vol. 14, no. 19, Art. no. 19, Jan. 2022, doi: 10.3390/rs14194687

[18] J. Drewes, S. Feder, and W. Einhäuser, “Gaze During Locomotion in Virtual Reality and the Real World,” Frontiers in Neuroscience, vol. 15, 2021, doi: 10.3389/fnins.2021.656913

[19] E. Suma, S. Finkelstein, M. Reid, S. Babu, A. Ulinski, and L. F. Hodges, “Evaluation of the Cognitive Effects of Travel Technique in Complex Real and Virtual Environments,” IEEE Trans. Visual. Comput. Graphics,

vol. 16, no. 4, pp. 690–702, Jul. 2010, doi: 10.1109/TVCG.2009.93

[20] A. Tikanmäki, T. Bedrník, R. Raveendran, and J. Röning, “The remote operation and environment reconstruction of outdoor mobile robots using virtual reality,” in 2017 IEEE International Conference on Mechatronics and Automation (ICMA), Aug. 2017, pp. 1526–1531. doi: 10.1109/ICMA.2017.8016043

[21] Y. Kato, “A remote navigation system for a simple tele-presence robot with virtual reality,” in 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015), Hamburg, Germany, Sep. 2015, pp. 4524–4529. doi: 10.1109/IROS.2015.7354020

[22] M. Hirschmanner, C. Tsiourti, T. Patten, and M. Vincze, “Virtual Reality Teleoperation of a Humanoid Robot Using Markerless Human Upper Body Pose Imitation,” in 2019 IEEE-RAS 19th International Conference on Humanoid Robots (Humanoids), Oct. 2019, pp. 259–265. doi: 10.1109/Humanoids43949.2019.9035064

[23] D. Whitney, E. Rosen, E. Phillips, G. Konidaris, and S. Tellex, “Comparing Robot Grasping Teleoperation Across Desktop and Virtual Reality with ROS Reality,” in Robotics Research, Cham, 2020, pp. 335–350. doi: 10.1007/978-3-030-28619-4_28

[24] J. Du, H. M. Do, and W. Sheng, “Human–Robot Collaborative Control in a Virtual-Reality-Based Telepresence System,” Int J of Soc Robotics, vol. 13, no. 6, pp. 1295–1306, Sep. 2021, doi: 10.1007/s12369-020-00718-w

[25] M. Wonsick and T. Padir, “A Systematic Review of Virtual Reality Interfaces for Controlling and Interacting with Robots,” Applied Sciences, vol. 10, no. 24, Art. no. 24, Jan. 2020, doi: 10.3390/app10249051

[26] A. L. Shelton and T. P. McNamara, “Orientation and Perspective Dependence in Route and Survey Learning,” Journal of Experimental Psychology: Learning, Memory, and Cognition, vol. 30, pp. 158–170, 2004, doi: 10.1037/0278-7393.30.1.158

[27] J. R. Brockmole and R. F. Wang, “Changing perspective within and across environments,” Cognition, vol. 87, no. 2, pp. B59-67, Mar. 2003, doi: 10.1016/s0010-0277(02)00231-7

[28] V. A. Diwadkar and T. P. McNamara, “Viewpoint dependence in scene recognition,” Psychological Science, vol. 8, pp. 302–307, 1997, doi: 10.1111/j.1467-9280.1997.tb00442.x

[29] F. Pazzaglia and H. A. Taylor, “Perspective, instruction, and cognitive style in spatial representation of a virtual environment,” Spatial Cognition and Computation, vol. 7, pp. 349–364, 2007, doi: 10.1080/13875860701663223

[30] P. W. Thorndyke and B. Hayes-Roth, “Differences in spatial knowledge acquired from maps and navigation,” Cognitive Psychology, vol. 14, no. 4, pp. 560–589, Oct. 1982, doi: 10.1016/0010-0285(82)90019-6

[31] L. F. Schneider and H. A. Taylor, “How do you get there from here? Mental representations of route descriptions,” Applied Cognitive Psychology, vol. 13, pp. 415–441, 1999, doi: 10.1002/(SICI)1099-0720(199910)13:5<415::AID-ACP602>3.0.CO;2-N

[32] H. A. Taylor and B. Tversky, “Spatial mental models derived from survey and route descriptions,” Journal of Memory and Language, vol. 31, no. 2, pp. 261–292, Apr. 1992, doi: 10.1016/0749-596X(92)90014-O

[33] H. A. Taylor, S. J. Naylor, and N. A. Chechile, “Goal-specific influences on the representation of spatial perspective,” Mem Cognit, vol. 27, no. 2, pp. 309–319, Mar. 1999, doi: 10.3758/bf03211414

[34] T. T. Brunyé, A. Gardony, C. R. Mahoney, and H. A. Taylor, “Going to town: Visualized perspectives and navigation through virtual environments,” Computers in Human Behavior, vol. 28, pp. 257–266, 2012, doi: 10.1016/j.chb.2011.09.008

[35] M. Al Zayer, P. MacNeilage, and E. Folmer, “Virtual Locomotion: A Survey,” IEEE Transactions on Visualization and Computer Graphics, vol. 26, no. 6, pp. 2315–2334, Jun. 2020, doi: 10.1109/TVCG.2018.2887379

[36] Z. Zheng, M. Pan, and W. Pan, “Virtual Prototyping-Based Path Planning of Unmanned Aerial Vehicles for Building Exterior Inspection,” in ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction, Oct. 2020, vol. 37, pp. 16–23. doi: 10.22260/ISARC2020/0004

[37] F. Argelaguet and M. Maignant, “GiAnt: stereoscopic-compliant multi-scale navigation in VEs,” Nov. 2016, pp. 269–277. doi: 10.1145/2993369.2993391

[38] P. Abtahi, M. Gonzalez-Franco, E. Ofek, and A. Steed, “I’m a Giant: Walking in Large Virtual Environments at High Speed Gains,” in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, New York, NY, USA, Mai 2019, pp. 1–13. doi: 10.1145/3290605.3300752

[39] S. Cmentowski, A. Krekhov, and J. Krüger, “Outstanding: A Multi-Perspective Travel Approach for Virtual Reality Games,” in CHI PLAY ’19: Proceedings of the Annual Symposium on Computer-Human Interaction in Play, Oct. 2019, pp. 287–299. doi: 10.1145/3311350.3347183

[40] L. Wang et al., “Occlusion Management in VR: A Comparative Study,” in 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, Mar. 2019, pp. 708–716. doi: 10.1109/VR.2019.8798025

[41] S. Chen et al., "UrbanRama: Navigating Cities in Virtual Reality," in IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 12, pp. 4685-4699, 1 Dec. 2022, doi: 10.1109/TVCG.2021.3099012

[42] J. J. LaViola, “A discussion of cybersickness in virtual environments,” SIGCHI Bull., vol. 32, no. 1, pp. 47–56, Jan. 2000, doi: 10.1145/333329.333344

[43] E. Chang, H.-T. Kim, and B. Yoo, “Virtual Reality Sickness: A Review of Causes and Measurements,” International Journal of Human-Computer Interaction, vol. 36, pp. 1–25, Jul. 2020, doi: 10.1080/10447318.2020.1778351

[44] M. E. McCauley and T. J. Sharkey, “Cybersickness: Perception of self-motion in virtual environments,” Presence: Teleoper. Virtual Environ., vol. 1, no. 3, pp. 311–318, Jan. 1992, doi: 10.1162/pres.1992.1.3.311

[45] C.-L. Liu and S.-T. Uang, “A study of sickness induced within a 3D virtual store and combated with fuzzy control in the elderly,” in 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery, May 2012, pp. 334–338. doi: 10.1109/FSKD.2012.6234149

[46] A. Krekhov, S. Cmentowski, K. Emmerich, M. Masuch, and J. Krüger, “GulliVR: A Walking-Oriented Technique for Navigation in Virtual Reality Games Based on Virtual Body Resizing,” in Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play, New York, NY, USA, Oktober 2018, pp. 243–256. doi: 10.1145/3242671.3242704

[47] “Unity 2021.1.20,” Unity, Jan. 03, 2021. https://unity3d.com/unity/whats-new/2021.1.20 (accessed Oct. 07, 2022)

[48] “Double Robotics - Telepresence Robot for the Hybrid Office.” https://www.doublerobotics.com/tech-specs.html (accessed Sep. 18, 2022)

[49] M. Hegarty, A. E. Richardson, D. R. Montello, K. Lovelace, and I. Subbiah, “Development of a self-report measure of environmental spatial ability,” Intelligence, vol. 30, no. 5, pp. 425–448, 2002, doi: 10.1016/S0160-2896(02)00116-2

[50] R. Kennedy, N. Lane, K. Berbaum, and M. Lilienthal, “Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness,” The International Journal of Aviation Psychology, vol. 3, pp. 203–220, Jul. 1993, doi: 10.1207/s15327108ijap0303_3

[51] D. A. Bowman, D. Koller, and L. F. Hodges, “Travel in immersive virtual environments: an evaluation of viewpoint motion control techniques,” in Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality, Albuquerque, NM, USA, 1997, pp. 45-52,. doi: 10.1109/VRAIS.1997.583043

[52] T. Hirzle, M. Cordts, E. Rukzio, J. Gugenheimer, and A. Bulling, “A Critical Assessment of the Use of SSQ as a Measure of General Discomfort in VR Head-Mounted Displays,” in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, New York, NY, USA, Mai 2021, pp. 1–14. doi: 10.1145/3411764.3445361

[53] P. Brown, P. Spronck, and W. Powell, “The simulator sickness questionnaire, and the erroneous zero baseline assumption,” Frontiers in Virtual Reality, vol. 3, 2022, doi: 10.3389/frvir.2022.945800

[54] R. Stoakley, M. J. Conway, and R. Pausch, “Virtual reality on a WIM: interactive worlds in miniature,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, USA, Mai 1995, pp. 265–272. doi: 10.1145/223904.223938

[55] C. A. Wingrave, Y. Haciahmetoglu, and D. A. Bowman, “Overcoming

World in Miniature Limitations by a Scaled and Scrolling WIM," in 3D User Interfaces (3DUI'06), Mar. 2006, pp. 11–16. doi: 10.1109/VR.2006.106

[56] D. Englmeier, W. Sajko, and A. Butz, "Spherical World in Miniature: Exploring the Tiny Planets Metaphor for Discrete Locomotion in Virtual Reality," in 2021 IEEE Virtual Reality and 3D User Interfaces (VR), Mar. 2021, pp. 345–352. doi: 10.1109/VR50410.2021.00057