

Unwinding Rotations Improves User Comfort with Immersive Telepresence Robots

Markku Suomalainen, Basak Sakcak, Adhi Widagdo, Juho Kalliokoski, Katherine J. Mimnaugh, Alexis P. Chambers, Timo Ojala and Steven M. LaValle
Center of Ubiquitous Computing, Faculty of Information Technology and Electrical Engineering,
University of Oulu, Oulu, Finland
firstname.surname@oulu.fi

Abstract—We propose unwinding the rotations experienced by the user of an immersive telepresence robot to improve comfort and reduce VR sickness of the user. By immersive telepresence we refer to a situation where a 360° camera on top of a mobile robot is streaming video and audio into a head-mounted display worn by a remote user possibly far away. Thus, it enables the user to be present at the robot’s location, look around by turning the head and communicate with people near the robot. By unwinding the rotations of the camera frame, the user’s viewpoint is not changed when the robot rotates. The user can change her viewpoint only by physically rotating in her local setting; as visual rotation without the corresponding vestibular stimulation is a major source of VR sickness, physical rotation by the user is expected to reduce VR sickness. We implemented unwinding the rotations for a simulated robot traversing a virtual environment and ran a user study (N=34) comparing unwinding rotations to user’s viewpoint turning when the robot turns. Our results show that the users found unwound rotations more preferable and comfortable and that it reduced their level of VR sickness. We also present further results about the users’ path integration capabilities, viewing directions, and subjective observations of the robot’s speed and distances to simulated people and objects.

Index Terms—telepresence; virtual reality; head-mounted display; VR sickness

I. INTRODUCTION

Immersive robotic telepresence refers to embodying a mobile robot in a far away location through a Head-Mounted Display (HMD). The robot is equipped with either a 360° camera or a regular camera with a fast pan-tilt-unit to follow the user’s view, streaming video to the HMD such that he can look around in the robot’s environment by turning his head (Fig. 1). This technology can greatly improve the experience for remote participants of *hybrid* events where a group meets locally, but others who would like to join cannot be there physically. Use of an HMD instead of a regular camera streaming into a regular screen as found in commercial telepresence robots (for example, GoBe and Double 3) has great potential to increase the *presence* of the remote user, and thus making them feel as if they really were there at the robot’s location.

However, the use of an HMD does not come without deficiencies, a major one being Virtual Reality (VR) sickness [1], [2] causing, for example, headache, blurred vision, and even vomiting. Different theories regarding the origin of VR

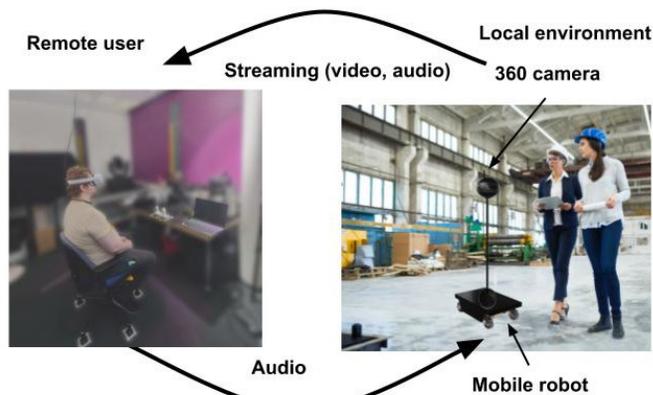


Fig. 1: Immersive robotic telepresence: a user wearing an HMD can embody a robot far away and join events and facility tours even if physically joining is impossible.

sickness are all related to *sensory mismatch* [3]; in immersive telepresence, a user’s eyes observe motion but the vestibular organ does not since the user does not move. Moreover, the increased immersion leading to an increased feeling of presence can cause users to feel uncomfortable about other aspects besides sickness, such as distance to walls or surprising turns [4], [5]. Thus, to make the benefits of immersive telepresence available for more people, care must be taken that the experience is comfortable and not sickening to users.

In this paper, we propose *unwinding the rotations* for the user immersed in a telepresence robot. Essentially, this means that when the robot rotates, the user’s viewpoint does not rotate along, such that the user always looks towards the same direction in the robot’s environment unless they rotate themselves (assuming that the users are either standing or sitting in a swivelling chair). With the known result of rotational motion causing more sickness than translational motion [6], [7], we target the rotational motions in an attempt to reduce VR sickness. We predict that, according to the sensory mismatch hypothesis [3], users will not get as sick when they have to perform the physical rotation themselves.

The main contributions of this paper are 1) the method for unwinding rotations, compensating for the rotational motion that the camera undergoes due to the robot motions, and 2) a user study showing that unwinding rotations reduces VR sickness, increases user comfort and is preferred. We performed the user study in simulation to eliminate possible confounding factors from a real robot causing additional

This work was supported by Business Finland (project HUMOR 3656/31/2019); Academy of Finland (projects PERCEPT 322637, CHIMP 342556); and the European Research Council (project ILLUSIVE 101020977)

sickness, such as vibrations or camera rotations around other axes. We further show that users’ path integration, the ability to mentally calculate updated position based on self-motions, does not deteriorate by unwinding the rotations. Additionally, we show that users rotate themselves in the chair to face the direction of motion, even though we did not prompt them to do it, and that they find it intuitive. Finally, we provide results from open-ended questions detailing reasons for preference of one method or the other, and update general recommendations regarding distances and speed of the robot in immersive telepresence. We believe that this paper makes an important contribution towards making immersive, HMD-based telepresence a more widely acceptable technology.

II. RELATED WORK

Telepresence, a term originally coined by Minsky [8], means being present in a remote location through technology. To increase the mobility of the user, it is natural to mount a camera and a screen on a mobile robot, often referred to as a telepresence robot. There is evidence that even this kind of simple telepresence robots increase the feeling of *presence*, the degree of “being there” often measured in VR research [9], of the remote user over a stationary video call [10]. Research has shown multiple uses for robotic telepresence, such as students who otherwise cannot attend lectures [11], [12], social interaction [13], and telemonitoring on the basis of medical consultation [14].

Despite showing promise, the increase in presence from having wheels may still not be enough to enable communication comparable to physical presence: Stoll et al. [15] showed that users communicating through a robot spoke significantly less and perceived tasks and communication more difficult than local, physical participants. This was very likely related to the lack of presence, since the feeling of presence is shown to make people act more naturally [16]; despite the increase due to the wheels, the immersion level of a flat screen was not enough to create sufficient presence¹. Combining the immersion of an HMD with wheels, i.e., using an immersive telepresence robot, has the potential to remedy this issue. Motivated by this potential, there has recently been an increasing amount of work about HMD-based telepresence. The immersive capacity of HMDs in telepresence has been shown beneficial in, for example, education [12], general communication [18] and even a real earthquake scenario [19], and there are recent advances on controlling such a robot [20]. However, an important aspect still remains: how to make the robot’s motions comfortable and non-sickening for the users.

VR sickness is one of the major risks preventing widespread use of HMDs. Whereas the umbrella term includes many issues causing users to feel sick, such as field of view and latency [21], in this paper we use the term to refer to Visually Induced Motion Sickness (VIMS) in VR context. It is estimated that only 10% of people have never experienced

VIMS symptoms in a vehicle [22]; as previous history in motion sickness is a good predictor for susceptibility to VR sickness [23], it is easy to grasp the gravity of the problem. In mainstream VR research, the problem is often alleviated by using teleportation or another locomotion model infeasible for a real robot. Beyond teleportation, other approaches to reduce VR sickness include slowing everything down (as optical flow correlates with experienced VR sickness [1]), using a rest frame, i.e., always keeping something constant in the user’s view [24], or narrowing the field of view [25]. However, all of these methods also deteriorate overall experience; slow motions are simply frustrating and a rest frame or narrower field of view reduce what the user sees and thus the feeling of presence [26]. Methods reducing VR sickness but not restricting the user’s field of view are needed to make immersive telepresence more available for everyone.

Besides VR sickness and user comfort, another metric of interest in this study is the path integration of the user, which should not deteriorate when unwinding rotations, since disorientation can cause anxiety and discontent [27]. Studies related to ours show that physical rotation increased spatial awareness (which correlates with path integration) over moving with a joystick [28] and a swivel chair increases spatial awareness over a fixed chair [29]. In this study we do not explicitly tell the users to rotate themselves along with the robot while watching the videos, but they do sit in a swivel chair making it easily possible. We hypothesized there to be no difference in path integration between unwinding rotations and automatic, coupled rotations, a sufficient result if confirmed since reducing VR sickness is the main goal of unwinding rotations. We believe, however, that our study will be an interesting addition to the research in path integration as well.

III. UNWINDING ROTATIONS

Unwinding rotations refers to rotating the camera frame to decouple the user’s viewpoint from the rotations of the robot. This is illustrated in Fig. 2: the robot is moving, either autonomously or controlled by a user directly, and will take a right at the corner. Suppose the user’s viewpoint before the turn is towards the wall at the end of the first corridor (view in Fig. 2a). If we are not unwinding the rotation that the camera frame undergoes (Fig. 2f) due to the robot motion, user’s viewpoint will change towards the robot heading as the robot rotates, even if she has not moved her head (Fig. 2c). The proposed unwinding rotations method ensures that, by rotating the camera frame (Fig. 2e), after the robot has made the turn, user’s viewpoint will no longer be towards the robot’s direction of motion but it will be towards the wall unless she physically moves her head (Fig. 2b).

The main requirement for rotations to be unwound is that the camera mounted on the robot is capable of adjusting the user’s viewpoint independently of the robot, through hardware (for example, a pan-tilt-unit) or software as in a 360° camera. Let \mathcal{R}_R be the set of all the rotations that can be applied to the robot to change its rotation, accordingly, let \mathcal{R}_R^{-1} denote the set corresponding to the inverses of all the elements in

¹We follow Mel Slater’s framework for terminology, in which immersion is the physical system’s capability of providing immersive experiences, and presence refers to the subjective feeling of being in a virtual location [17]

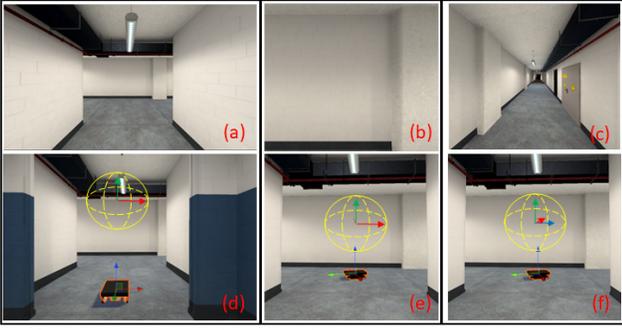


Fig. 2: The concept of unwinding rotations, assuming that user did not move her head. (a) Initial view of the immersed user. (b) View after the rotation is unwound. (c) View after the rotation is not unwound. (d) Initial robot configuration with robot and camera (in yellow sphere) coordinate frames. Robot and camera coordinate frames corresponding to (e) unwound and (f) not unwound rotation.

\mathcal{R}_R . Note that, by the definition of the rotation matrix, the inverse always exists. Then, for unwinding rotations to be possible, the set of rotational transformations that can be applied to the camera frame \mathcal{R}_C should contain \mathcal{R}_R^{-1} , i.e., $\mathcal{R}_R^{-1} \subseteq \mathcal{R}_C$. Essentially, this means that the camera frame should be capable of rotating to compensate for the robot rotation. For example, a 360° camera can unwind any rotation through software, but a pan-tilt unit that cannot roll cannot be used to unwind rotations on a drone rotating about all three axes; however, such a pan-tilt-unit can be used to unwind all the rotations of a telepresence robot traversing a flat surface. Note that very small rotations can also be compensated using, for example, the post-rendering image warp [30].

Consider a robot carrying a camera and moving in 3D space (for example, a drone or the end effector of a manipulator) such that $\mathcal{R}_R \in SO(3)$. Let $R_{R,k} \in \mathcal{R}_R$ be the rotation matrix corresponding to the robot orientation at time step $k \in \mathbb{Z}_{\geq 0}$. We assume that the robot orientation, hence $R_{R,k}$, is known or can be estimated accurately at k , guaranteed for robotic systems which have sufficient sensing and filtering capabilities. Then, we can define unwinding rotations as rotating the camera frame such that any point $p_c \in \mathbb{R}^3$ represented in the camera frame is related to the point p'_c , whose coordinates are expressed in the rotated camera frame, through

$$p'_c = R_{C,k} p_c \quad (1)$$

in which $R_{C,k} = R_{R,k}^{-1} \in \mathcal{R}_C$. This operation negates the rotation of the viewpoint caused by the robot rotation. Thus, the users need to rotate themselves in order to face the direction that the robot is or face the same direction in the virtual environment for the whole motion.

The study presented in this paper considers a mobile robot moving in a two-dimensional plane, meaning that $R_{R,k}$ corresponding to the robot orientation θ_k at time step k describes a rotation about the axis orthogonal to the plane that the robot is moving in. Therefore, for this special case, \mathcal{R}_R is a subspace of $SO(3)$. In particular, if we consider that the axis about which the robot rotates is the z -axis (pointing up), then

$$R_{R,k} = \begin{bmatrix} R(\theta_k) & 0_{2 \times 1} \\ 0_{1 \times 2} & 1 \end{bmatrix} \quad (2)$$

in which $R(\theta_k) \in SO(2)$ is the rotation in the x - y plane. Then, we can use (1) to unwind the rotations applied to the camera frame due the changes in the robot orientation.

IV. HYPOTHESES

We pre-registered the following four hypotheses, together with the procedure and analyses to be used in the study, in Open Science Foundation (OSF) <https://osf.io/eks6t>. From now on, we will refer to the **unwound rotations condition as UR**, for which the user's viewpoint does not change unless she physically rotates her head. Accordingly, we will refer to **Coupled rotations as CR**, for which the user's viewpoint always rotates when the robot rotates.

- H1:** Less VR sickness in UR condition as indicated by lower total weighted Simulator Sickness Questionnaire (SSQ) score.
- H2:** The UR condition is more comfortable as indicated by asking directly which was more comfortable (forced-choice).
- H3:** The UR condition is preferred as indicated by asking directly which the user preferred (forced choice).
- H4:** No difference in path integration across conditions when measured by error in angle when asked to point towards the point of origin after the robot has finished the path.

H1 is based on the literature; as explained in Section II, the sensory conflict theory postulates that a large part of VR sickness is caused by a mismatch between vision and vestibular organs. Therefore, we predicted that if the user experiences only the rotations that he performs physically, there will no sensory conflict; as a result, sickness should not be induced. **H2** and **H3** were then prompted by the assumption that people do not want to be sick, and we did not foresee other major downsides that could make UR not preferred or uncomfortable. **H4** is derived from the literature as well; there seemed to be conflicting results regarding path integration, as presented in Section II, which prompted us to propose no difference. The reason we test this is to show that UR does not deteriorate this important aspect of the experience.

V. EXPERIMENTS

A. Study setup

The image on the left in Fig. 1 shows the setup used in this study. During the experiment, the participants sat on a swivel chair without wheels, to be able to easily rotate themselves as much as they wanted without translations possibly caused by wheels. We chose sitting over standing for a few reasons; first, many people seem to prefer sitting over standing in VR [31]; second, it is more inclusive since standing is physically more difficult for many people, and third; this way we can be sure that the people do not translate (take steps around), which could confound the study. To avoid the participants getting entangled with the cable, we installed a pulley system that feeds the cable from the ceiling.

The virtual environment was designed using Blender and Unity. It is a combination of a real room at the University of Oulu and custom designed, imaginary areas (see Fig. 3 for

screenshots from the environment); thus, even the people who are familiar with the university have no advantage in the path integration task. The virtual robot has been modelled after a real robot owned by the research group, a balanced differential drive robot with the 2 driving wheels in the middle and 4 caster wheels in corners to keep it steady (see Fig. 2). We ensured that the dynamics governing the robot motion in simulation is sufficiently realistic. A virtual 360° camera is attached 1.5 meters above the base, from which the user sees the virtual world; this is a height suggested by [32] for 360° videos. To increase the ecological validity of the environment and the path, there are simulated people walking in the environment (Fig. 3a) that the robot occasionally evades.

We used Nav2 project of Robot Operating System (ROS) 2 Galactic for the navigation of the robot, and Unity as the physics-based simulator. The robot route is described as a sequence of waypoints, consequently, the path is calculated and tracked autonomously. In particular, we chose Theta* [33] as the path planning algorithm and DWB (a controller based on Dynamic Window Approach [34]) as the controller to track the computed path. Since navigation is probabilistic in nature, due mainly to the localization and control methods used, one instance of the robot motion, together with the motion of the simulated humans, was recorded and then played back to the users in a Unity executable to avoid small differences in emerging motions. In the recorded motion, the robot moves at a maximum speed of 1 m/s and rotates with velocities within the range $[0, 1]$ rad/s with maximum rotational acceleration 3.2 rad/s^2 as suggested in [4], [5]. The path is 129.5 meters long, the video lasts for 2 minutes 18 seconds, and the total amount of rotation that the robot executes during the video is 438°. The lowest distance to a wall or another inanimate object was 0.9m and the closest distance to people passing by 1 m; these values correspond to the distance accepted by users in [4]. The application was viewed through Oculus Quest 2 via a link cable, with refresh rate 80Hz and resolution 5152x2608.

B. Procedure

The two videos of the UR and CR conditions were presented to the participants in a counterbalanced order such that both videos were seen first and second equal number of times by both men and women. Upon arrival, participants were greeted by a researcher and asked for consent, after which the experimenter asked the participants if they were feeling nauseous or had a headache in an effort to pre-screen people feeling sick already before the experiment (we would have re-scheduled a sick-feeling participant). Then, the users were asked to move to the swivel chair and adjust the chair height such that they could easily rotate around. Next, the experimenter read out the instructions, told the participant how to put on the HMD and asked them to rotate around once more while wearing the HMD; finally, the experimenter asked them to orient themselves towards “the corridor with the fire hose”, which was the direction the robot would start moving. Once this was confirmed and done (the experimenter could always see the participants’ view on a mirrored screen) the

experimenter started the video. During the video, the head orientation of the participant was recorded.

When the video finished, the participants were asked not to take off the HMD but instead grab the controllers from a desk in front of them. Then, the users were asked to point towards the beginning of the robot’s path, which was clarified as the place with the fire hose to make sure participants understood this correctly. Then, the participants were asked to remove the HMD and fill out a questionnaire regarding their experience, after which the same procedure was repeated for the second video. At the end, the participants were rewarded with 20€ Amazon vouchers. Finally, we took the recommended precautions regarding Covid-19 when running the study. The experimenter was always wearing a mask and kept distance from the participant except if help was needed with the HMD. With the HMD disposable face covers were used.

C. Measures

The path integration error was measured at the robot’s final position as the absolute angle between the direction indicated by the participant and the direction of the origin (robot start position). Thus, the error lies within the interval $[0°, 180°]$. We computed the direction using the projection of the line segment protruding from the controller (see Fig. 3c) onto the plane that the robot moves in.

Head motion data was used to compute the average deviation from the robot heading to see whether participants aligned themselves with the direction of motion as they would if the rotations were not unwound. Deviation was calculated as the absolute angle difference between the robot orientation and the head orientation in the plane that the robot moves in; absolute angle was used to avoid bias due to the robot’s turns not being equally divided between left and right. Furthermore, head orientation is plotted along the robot path to qualitatively analyze the emergent view patterns across conditions.

In the questionnaires, we measured sickness with the SSQ [35], an established questionnaire for measuring sickness in VR by presenting 16 possible sickness symptoms, which the participants gauge on a scale none (0) to severe (3). The answers are weighted for a maximum score of 236 [36]. Additionally, we used 7-point Likert-scale questions, forced-choice questions comparing the two videos, and open-ended questions about reasons for some choices and demographic questions. The first question after each video was about the user’s confidence in the path integration, after which SSQ was administered and we asked how comfortable the experience was on Likert-scale. After the second video, these questions were followed by forced-choice questions of choosing the preferred, more comfortable and more intuitive condition for the participant. Then, we asked Likert-scale questions about how the participants felt regarding distance to walls, distance to humans, and the linear speed of the robot; these questions were asked to make further suggestions for researchers working on autonomous motion planners for telepresence robots. Finally, we asked about VR and gaming experience and demographics. The whole questionnaire can be found in Appendix 1.

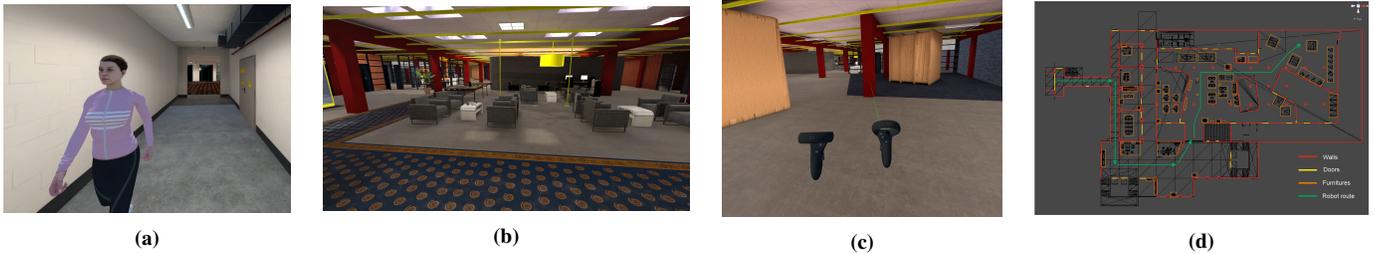


Fig. 3: Screenshots from the environment from the user’s view; (a) a person passing by in a corridor; (b) general view; (c) raycasting with controllers at the end to point towards origin; (d) a birds-eye view of the environment with the taken path illustrated.

D. Participants

Based on the effect size from a power analysis of a previous study comparing two immersive telepresence experiences [37], the necessary sample size was determined as 32 participants. We required an exact split between male and female participants due to earlier, contradicting evidence that gender may have an effect on the susceptibility of VR sickness, either women suffering more [38] or less [39] from VR sickness. When two participants from the first 32 recruited preferred not to report their gender, we recruited two more participants to be certain of the gender balance, ending up with 34 participants recruited from the university of Oulu campus and the surroundings. We used only the gender-specified 32 participants for testing for VR sickness differences between genders, and all 34 participants for all other analyses. Participants’ age ranged from 18 to 48 with the mean at 26.21 ($sd = 7.1$). All participants reported having normal or corrected-to-normal vision and none of the participants were colorblind. The responses of the participants to how often they use VR systems were: 41.2% never used any before, 35.3% just a couple of times and at least once, 11.8% once or twice a year, 3% once or twice a month, and 8.8% several times a week. Their responses to how often they play computer games were: 11.8% never played before, 8.8% once or just a couple of times ever, 20.6% once or twice a year, 11.8% once or twice a month, 17.6% once or twice a week, 14.7% several times a week, and 14.7% every day.

VI. RESULTS

Four confirmatory hypothesis tests were preregistered, as well as several topics for exploratory analysis. All tests were run in SPSS with (two-tailed) significance levels set to 0.05 and with a 95% confidence interval.

A. Confirmatory Results

A Wilcoxon Signed-Ranks test (two sided) is performed to compare the differences between the total weighted SSQ scores for UR ($Mdn = 9.35$) and CR ($Mdn = 16.83$) conditions (see Fig. 4a for the score distributions). The test indicated that Unwound Rotations (UR) elicited significantly lower SSQ scores compared to Coupled Rotations (CR), $Z = -3.46$, $p = .001$, $r = 0.59$.

Fig. 5 shows the distributions of the responses given by the participants to the forced-choice questions regarding preference and comfort. When asked “Which of the two videos do

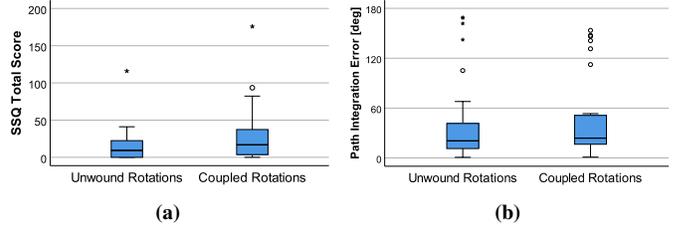


Fig. 4: Comparison of (a) total weighted SSQ scores, and (b) path integration errors.

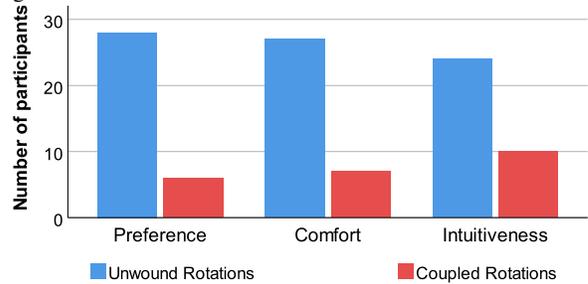


Fig. 5: The distributions of responses to the questions regarding preference, comfort, and intuitiveness.

you prefer?” 28 out of 34 participants (82.4%) selected the UR condition. An exact binomial test with exact Clopper-Pearson 95% CI indicated that this tendency in preference is statistically significant, $p = .000$ (two-sided, one-sided $p = .000$) and had a 95% CI of 65.5% to 93.2%. Consequently, 27 participants (79.4%) selected the UR condition in response to the question “Which of the two videos is more comfortable?”. An exact binomial test with exact Clopper-Pearson 95% CI was performed, showing that the condition UR was found significantly more comfortable, $p = .001$ (two-sided, one-sided $p = .0005$) and had a 95% CI of 62.1% to 91.3%.

The medians of the path integration error distributions for UR and CR conditions were 20.57°, and 23.9°, respectively (see Fig. 4b for the error distributions). A Wilcoxon Signed-Ranks test (two sided) indicated that UR did not elicit any significant change in the path integration errors, $Z = -0.71$, $p = .478$, $r = 0.12$.

B. Exploratory Results

In addition to the confirmatory analyses we performed exploratory analyses to interpret the results better.

1) *Quantitative Data:* We checked if there were any carryover effects on sickness with 5-10 minute breaks and did not observe an order effect in the weighted total SSQ scores across conditions UR and CR, as indicated by a Wilcoxon

Signed-Ranks test (two-sided), $Z = 1.1$, $p = .272$, $r = 0.19$. We also looked for the potential effect of gender on sickness. For this analysis, we used only the gender-specified participants ($n = 32$), balanced equally between men and women. For CR condition, we did not observe any significant difference between the total weighted SSQ scores of women ($Mdn = 33.66$) compared to men ($Mdn = 14.96$), as indicated by a Mann-Whitney U test (two-sided), $U = 76.5$, $Z = -1.96$, $p = .051$, $r = 0.35$. Similarly, there was not a statistically significant difference between the SSQ scores of men ($Mdn = 3.74$) and women ($Mdn = 13.09$) in UR condition either, as indicated by a Mann-Whitney U test (two-sided), $U = 98.5$, $Z = -1.14$, $p = .26$, $r = 0.2$.

In addition to the forced-choice question, Likert-scale is used to measure the user comfort. Comparing the comfort ratings in UR condition ($Mean = 5.94$) with the ones in CR condition ($Mean = 5.54$), we found that UR elicited a statistically significant increase in the comfort rankings, as indicated by a Wilcoxon Signed-Ranks test (two sided), $Z = 2.16$, $p = .032$, $r = 0.37$.

We observed an order effect in the path integration errors. Watching the same environment twice resulted in a statistically significant decrease in the errors after the second video compared to the first one, as indicated by a Wilcoxon Signed-Ranks test (two-sided), $Z = -2.11$, $p = .035$, $r = 0.36$. However, we remind the reader that the experiment was counterbalanced such that half of the participants had UR as the starting condition and the remaining half had CR to keep potential order effects equal in both conditions.

The head motion data is analyzed to see whether people aligned themselves with the robot heading in UR condition. The distributions of average deviations from the robot heading for both conditions, UR and CR, followed a normal distribution as indicated by a Shapiro-Wilk test, $W(34) = 0.94$, $p = .055$, $W(34) = 0.97$, $p = .45$, respectively. Therefore, a paired t-test was run to determine whether there was a statistically significant mean difference between the average deviations in two conditions. The mean average deviation was higher in UR condition ($32.14^\circ \pm 13.16^\circ$) as opposed to in CR condition ($27.53^\circ \pm 12.46^\circ$); a statistically significant increase of 4.6° (95% CI, 0.76° to 8.43°), $t(33) = 2.44$, $p = .02$, $d = 0.42$.

We tested whether one condition felt more intuitive for the participants by explicitly asking: Which of the two experiences feels more intuitive for you? 24 participants (71%) felt that the condition UR felt more intuitive (Fig. 5). An exact binomial test with exact Clopper-Pearson 95% CI indicated that this tendency to pick UR condition was significant and had a 95% CI of 52.5% to 84.9%, $p = .024$ (two-sided).

We also asked Likert-scale questions about how the users perceived the speed, distance to walls and distance to people while moving with the robot. The statistics can be seen in Table I; as mentioned in Section V-A, the values for distances and speed were mainly based on [4]. The distance to walls and speed were considered, on average, suitable (median 4 on 7-point Likert scale), whereas the distance to people was on average considered slightly too short (median 3).

	Value	Median	Variance
Distance to walls	min. 0,9 m	4	0.78
Distance to people	min. 1 m	3	0.86
Speed	max. 1 m/s	4	0.94

TABLE I: The values of certain attributes of the path, and the median and variance of 7-point Likert scale subjective opinions (1 meaning too small/too close).

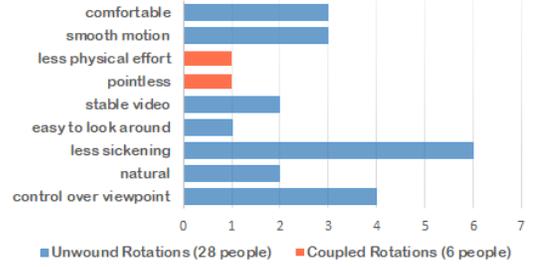


Fig. 6: Frequently found codes for question “Please explain why you prefer that video.”.

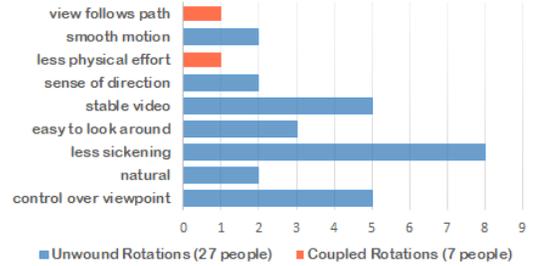


Fig. 7: The frequently found codes from the question “Please explain why that video is more comfortable.” which was asked after the participant watched both videos.

2) *Qualitative data:* The open-ended data was analyzed using the thematic analysis method with inductive approach [40]. In the first stage of analysis, two researchers independently identified codes from the response data and mutually agreed on the codings in the second phase. It was not mandatory for participants to answer the open-ended questions, and these fields were often left blank.

The frequent codes used in open-ended question asking why participants preferred the video can be seen in Fig. 6, divided by which video was preferred. The greatest number of comments (6 comments given by 21% of the 28 participants who preferred UR) said that UR was less sickening (for example, “*Because I have to turn around by myself, so it will reduce the vertigo feeling.*”). The next most frequently mentioned keywords were control over viewpoint (4 (14%), “*My head orientation was not forced*”), comfort (3 (11%), “*I think second video is more comfortable for eyes when the scene turning*”) and smoothness of the motion (3 (11%), “*If I try to remember, there might have been smoother movement in the second video.*”). In contrast, one person preferred CR because the physical effort of rotating in the chair felt pointless.

The responses to open-ended questions regarding choosing one video as more comfortable are seen in Fig. 7. Here, less sickening was even more evidently the biggest factor (8 (30%)), with control over viewpoint being the second (5 (19%)) with similar arguments as in preference. Codewords frequently not found in the question for preference but found

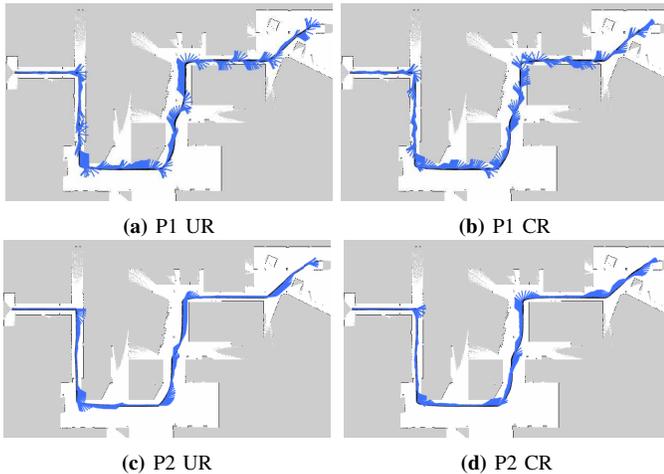


Fig. 8: Viewpoints of participants P1 and P2 along the robot's path.

here were stable video (5 (19%), “*The second one was not stable, the view kept moving, was terrible.*”) and easiness to look around (3 (11%), “*It was easier and more natural to look around and I experienced no motion sickness.*”).

VII. DISCUSSION

A. Original hypotheses

The results confirmed all hypotheses **H1-H4**. We were confident that **H1** would hold due to evidence from literature. However, whether less sickness would be enough to confirm **H2** and **H3** was an open question, since the immersion of an HMD and the need to physically rotate to look towards the direction of motion could have caused unexpected adversarial effects. Eventually, there was only one comment in the open-ended questions about not enjoying turning in the chair (“*Since I have no control on the direction, I don't want to have to turn the chair towards the direction of the video.*”), but, as seen in Figs. 6 and 7, more people enjoyed being in control of the viewpoint and did not mind having to rotate the chair (“*The first one felt much smoother and more comfortable. I did not mind having to turn around in my chair to look to the direction I was going.*”, “*I preferred the lack of forced turns.*”).

We predicted, as stated in **H4**, that UR does not deteriorate the spatial awareness of the users, and the results confirmed it. Even though there is some evidence in literature that physical rotations increase spatial awareness [28], we did not explicitly tell the participants to always turn to face the direction of motion of the robot. Thus, if the participants decided not to follow up with the robot's motion, it could have made path integration more challenging. However, the increase in the mean of average deviation from the robot heading in UR as compared to CR was less than 10° . This is negligible considering the total amount of rotation that the robot undergoes (438°) and the corresponding compensatory motions the user needs to do. Thus, we conclude that people aligned themselves with the robot as they would if the rotations were not unwound. This is supported by the qualitative analysis of the plots that show viewpoints along the robot path (see Fig. 8). The users were looking forwards approximately equally in both conditions

regardless of how much they looked around, prompting us to deduce that the users rotated to face the robot's direction of motion. Moreover, both the score on intuitiveness being in favor of UR, as well as some open-ended answers regarding intuitiveness (“*In the first video from what I can recall, I barely moved the chair. Nonetheless, on the second video I did move it every time the video changed directions, even though the instructors did not prompt me to do so.*”) strengthened our belief that users do find it intuitive and natural to rotate along with the robot to help confirming **H4**.

B. Exploratory data

We also wanted to contribute to the literature in terms of the recommended values for the clearance and speed of an immersive telepresence robot, since there is little research on the topic. The closest distance to walls in this study ($0.9m$) was significantly larger than the one suggested in [4] ($> 0.4 m$) for an immersive telepresence situation. Even though there were some comments like “*On some moments it felt too close. For instance, in the first turn, it felt as If I was going to crash.*”, the median rating was 4 (not too close or not too far). This suggests that $0.9 m$ was an appropriate distance to keep from inanimate objects and walls. Interestingly, despite the slightly larger distance from the (simulated) people ($1 m$), participants considered that the robot passed slightly too close ($Mdn = 3$). This is reflected also in their responses to the open-ended questions (for example, “*For some of the people crossing, it felt as If I should step a bit away, as I perceived them a bit too close as well. But overall it was okay.*” and “*felt like I was going to run into the woman in the hallway*”). Based on the results in this study we recommend that a minimum of $1 m$ distance should be kept from people, but further research is needed to determine an appropriate value. Whereas there is research on human-robot proxemics in VR [41], it lacks the telepresence perspective. Finally, most of the participants rated the robot speed of $1 m/s$ as suitable ($Mdn = 4$) with several answers to the open-ended question similar to “*It felt like a walking speed of a normal person.*”. On the other hand, some people perceived it to be slightly too slow (“*It was slightly too slow, but not terrible.*”) or too fast (“*It could be a bit slower so we could see the environment with more details*”). Enabling the user to adjust the speed, or having an adaptive speed controller [42] can remedy the discrepancy in the perceived speed among users (as supported by higher variance).

We did not observe carryover effects in the SSQ scores. In literature, a large variation in recovery times has been observed, with recovery times shorter than 10 minutes risking carryover effects [43]; in this study the breaks between videos were approximately 5-10min. One possible reason for us not observing carryover effects, regardless of the short and uncontrolled recovery time, is that the sessions were short with the 2min 18s duration; in many publications, the sessions are longer, and there are suggestions that even 55-70 minute sessions times would be acceptable to avoid excessive VR sickness [44]. Even if there were carryover effects, the counterbalancing would counter for them.

Until recently, women were considered to suffer more from VR sickness compared to men [38]. However, a recent study proposes that the design of commercial HMDs, which are based on the interpupillary distance of men, can cause women to suffer more [39]. Despite observing no significant differences in total weighted SSQ scores across genders, we found that the median of scores corresponding to women were higher as compared to men in both conditions. In particular, this increase was close to being borderline significant for the CR condition. This result was not surprising and can hopefully help designing future studies and better HMDs.

In the open-ended answers for both comfort and preference there were six comments (the same person for both questions) justifying their choice by the order of the videos, such as *"I was more familiar with the setting since I had seen it before."* and *"It was more exciting because it was the first time for me to wear vr-glasses and the second one was more like repeating that experience"*. The choices for both questions were four people choosing UR and two people choosing CR; this shows that the counterbalancing worked and the order effect did not corrupt the results. Evidently, using the same environment and the same robot path to avoid confounding factors lead to directing participants' attention to aspects that were tangential to the study. Even though counterbalancing ensured that this occurred approximately evenly for both conditions, injecting small changes to the environment in between conditions can be considered for future studies.

C. Limitations and Future Work

The observed order effect in the path integration could have been avoided by telling participants before the first video exactly what would be asked from them afterwards. The participants were not told their error after the first video, but simply focusing more on the second video caused the observed order effect. Nonetheless, due to counterbalancing there is no reason to doubt the overall result of not having a significant difference in the errors. Thus, it seems that path integration mainly depends on how much the participants focus on it, and not as much on physical or automatic, coupled rotations.

An immediate future work is to test the proposed method in real world experiments using a system comprising a wheeled robot equipped with a 360° camera, and an Inertial Measurement Unit (IMU) to estimate the rotations that are applied to the camera frame. Although there is some evidence that 360° videos induce more sickness than virtual environments [45], the physiological causes of VR sickness remain the same. Thus, we expect unwinding rotations to improve the experience, despite the likelihood of having more vibrations even with filtering. Further future work involves applying the method proposed in this paper to robots that are capable of rotating about more axes than just one, such as a drone, and researching whether the method can be applied to manually controlled vehicles as well.

We decided not to administer the participants the Slater-Usoh-Stetted (SUS) [46], [47] presence questionnaire, since we did not expect a significant difference in presence between

conditions. However, there are some hints in the open-ended questions pointing towards increased presence with UR, such as *"I felt more immersed and I didn't feel the changes of direction as I felt in the second video."* and *"It feels like I was more in control of the experience than in the second video. It felt like I was more present and could effect the experience better."* (we note that "immersion" is often understood as what we defined as "presence"). One possible reason why this could affect presence is that when users rotate more to match with the robot's direction of motion, the *sensorimotor contingencies*, meaning that the user's physical actions are matched visually in the virtual world, become more utilized; for example, one user mentioned that *"Because I could turn around in my chair physically, I could adjust easily where I was looking. Of course I could turn around in my chair during the second video as well, but it felt futile to do so."* Sensorimotor contingencies have been shown to be an important building block of presence [48]. An interesting follow-up study would be to compare the presented approach to an automatically turning chair [49]; in theory the suffered VR sickness should be the same, but we may observe a difference in presence.

The general concept of immersive telepresence requires further research on many fronts to become competitive against current commercial flat-screen alternatives, especially for tasks that require bidirectional communication. It is simple to show the remote user's face on a screen carried by the robot using standard solutions. However, immersion brings two issues: first, if a 360° camera is used, should the people around the robot know where the remote user is looking at? Even though we are not aware of a research confirming this, the likely answer is yes. In this case, research should be done to reveal, for example, whether a simple solution of using an array of lights indicating the remote user's view direction would suffice. Alternatively, a rotating screen could be attached to the robot, showing the remote user's face in the direction that he is looking at. However, this brings up the second problem; when the user is wearing an HMD, the user's face cannot be shown as simply. Fortunately, there is research towards rendering the face of an HMD-user on a screen as if she is not wearing an HMD [50]. As these technologies would likely ease two-way communication with immersive telepresence robots, we believe they will gain in popularity in the near future.

VIII. CONCLUSION

We show that unwinding the rotations decreases VR sickness, increases user comfort, and is preferred by the users when immersed in an autonomously moving telepresence robot. Further, users physically rotate themselves to face the direction of motion even without explicit instructions to do so and find it an intuitive way to embody an autonomously moving robot. We also observed no differences in path integration abilities between unwinding rotations and letting the user's viewpoint rotate. Thus, unwinding rotations shows great potential to make immersive telepresence more popular.

REFERENCES

- [1] J. J. LaViola Jr, "A discussion of cybersickness in virtual environments," *ACM Sigchi Bulletin*, vol. 32, no. 1, pp. 47–56, 2000.
- [2] S. M. LaValle, *Virtual reality*. Cambridge University Press, 2021.
- [3] J. T. Reason and J. J. Brand, *Motion sickness*. Academic press, 1975.
- [4] K. J. Mimnaugh, M. Suomalainen, I. Becerra, E. Lozano, R. Murrieta-Cid, and S. M. LaValle, "Analysis of user preferences for robot motions in immersive telepresence," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2021.
- [5] M. Suomalainen, K. J. Mimnaugh, I. Becerra, E. Lozano, R. Murrieta-Cid, and S. M. LaValle, "Comfort and sickness while virtually aboard an autonomous telepresence robot," in *EuroXR 2021: Virtual Reality and Mixed Reality*. Cham: Springer International Publishing, 2021, pp. 3–24.
- [6] A. Kemeny, P. George, F. Mérienne, and F. Colombet, "New VR navigation techniques to reduce cybersickness," *Electronic Imaging*, vol. 2017, no. 3, pp. 48–53, 2017.
- [7] S. Hu, K. A. McChesney, K. A. Player, A. M. Bahl, J. B. Buchanan, and J. E. Scozzafava, "Systematic investigation of physiological correlates of motion sickness induced by viewing an optokinetic rotating drum." *Aviation, space, and environmental medicine*, 1999.
- [8] M. Minsky, "Telepresence," 1980.
- [9] R. Skarbez, F. P. Brooks, Jr, and M. C. Whitton, "A survey of presence and related concepts," *ACM Computing Surveys (CSUR)*, vol. 50, no. 6, pp. 1–39, 2017.
- [10] I. Rae, B. Mutlu, and L. Takayama, "Bodies in motion: mobility, presence, and task awareness in telepresence," in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2014, pp. 2153–2162.
- [11] M. Rueben, M. Syed, E. London, M. Camarena, E. Shin, Y. Zhang, T. S. Wang, T. R. Groechel, R. Lee, and M. J. Matarić, "Long-term, in-the-wild study of feedback about speech intelligibility for k-12 students attending class via a telepresence robot," *arXiv preprint arXiv:2108.10456*, 2021.
- [12] J. Botev and F. J. Rodríguez Lera, "Immersive robotic telepresence for remote educational scenarios," *Sustainability*, vol. 13, no. 9, p. 4717, 2021.
- [13] A. Kristoffersson, S. Coradeschi, and A. Loutfi, "A review of mobile robotic telepresence," *Advances in Human-Computer Interaction*, vol. 2013, 2013.
- [14] K. A. R. Carranza, N. J. B. Day, L. M. S. Lin, A. R. Ponce, W. R. O. Reyes, A. C. Abad, and R. G. Baldovino, "Akibot: a telepresence robot for medical teleconsultation," in *2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*. IEEE, 2018, pp. 1–4.
- [15] B. Stoll, S. Reig, L. He, I. Kaplan, M. F. Jung, and S. R. Fussell, "Wait, can you move the robot? examining telepresence robot use in collaborative teams," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 14–22.
- [16] M. V. Sanchez-Vives and M. Slater, "From presence to consciousness through virtual reality," *Nature Reviews Neuroscience*, vol. 6, no. 4, pp. 332–339, 2005.
- [17] M. Slater and S. Wilbur, "A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments," *Presence: Teleoperators & Virtual Environments*, vol. 6, no. 6, pp. 603–616, 1997.
- [18] J. Du, H. M. Do, and W. Sheng, "Human-robot collaborative control in a virtual-reality-based telepresence system," *International Journal of Social Robotics*, pp. 1–12, 2020.
- [19] F. Negrello, A. Settini, D. Caporale, G. Lentini, M. Poggiani, D. Kanoulas, L. Muratore, E. Luberto, G. Santaera, L. Ciarleglio *et al.*, "Humanoids at work: The walk-man robot in a postearthquake scenario," *IEEE Robotics & Automation Magazine*, vol. 25, no. 3, pp. 8–22, 2018.
- [20] G. Baker, T. Bridgwater, P. Bremner, and M. Giuliani, "Towards an immersive user interface for waypoint navigation of a mobile robot," in *The Second International Workshop on Virtual, Augmented and Mixed Reality for Human-Robot Interaction*, 2020.
- [21] 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, no. 17, p. 1–25, 2020.
- [22] B. D. Lawson *et al.*, "Motion sickness symptomatology and origins," 2014.
- [23] J. F. Golding, "Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness," *Brain research bulletin*, vol. 47, no. 5, pp. 507–516, 1998.
- [24] Z. Cao, J. Jerald, and R. Kopper, "Visually-induced motion sickness reduction via static and dynamic rest frames," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2018, pp. 105–112.
- [25] J. Teixeira and S. Palmisano, "Effects of dynamic field-of-view restriction on cybersickness and presence in hmd-based virtual reality," *Virtual Reality*, pp. 1–13, 2020.
- [26] S. Weech, S. Kenny, and M. Barnett-Cowan, "Presence and cybersickness in virtual reality are negatively related: a review," *Frontiers in psychology*, vol. 10, p. 158, 2019.
- [27] R. P. Darken and B. Peterson, "Spatial orientation, wayfinding, and representation," in *Handbook of virtual environments*. CRC Press, 2002, pp. 533–558.
- [28] B. E. Riecke, B. Bodenheimer, T. P. McNamara, B. Williams, P. Peng, and D. Feuereissen, "Do we need to walk for effective virtual reality navigation? physical rotations alone may suffice," in *International Conference on Spatial Cognition*. Springer, 2010, pp. 234–247.
- [29] Y. Hong, A. MacQuarrie, and A. Steed, "The effect of chair type on users' viewing experience for 360-degree video," in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, 2018, pp. 1–11.
- [30] W. R. Mark, L. McMillan, and G. Bishop, "Post-rendering 3d warping," in *Proceedings of the 1997 symposium on Interactive 3D graphics*, 1997, pp. 7–ff.
- [31] D. Zielasko and B. E. Riecke, "To sit or not to sit in vr: Analyzing influences and (dis) advantages of posture and embodied interaction," *Computers*, vol. 10, no. 6, p. 73, 2021.
- [32] T. Keskinen, V. Mäkelä, P. Kallioniemi, J. Hakulinen, J. Karhu, K. Ronkainen, J. Mäkelä, and M. Turunen, "The effect of camera height, actor behavior, and viewer position on the user experience of 360 videos," in *2019 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, 2019, pp. 423–430.
- [33] A. Nash, S. Koenig, and C. Tovey, "Lazy theta*: Any-angle path planning and path length analysis in 3d," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 24, no. 1, 2010.
- [34] D. Fox, W. Burgard, and S. Thrun, "The dynamic window approach to collision avoidance," *IEEE Robotics & Automation Magazine*, vol. 4, no. 1, pp. 23–33, 1997.
- [35] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.
- [36] P. Bimberg, T. Weissker, and A. Kulik, "On the usage of the simulator sickness questionnaire for virtual reality research," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 2020, pp. 464–467.
- [37] I. Becerra, M. Suomalainen, E. Lozano, K. J. Mimnaugh, R. Murrieta-Cid, and S. M. LaValle, "Human perception-optimized planning for comfortable vr-based telepresence," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6489–6496, 2020.
- [38] L. Rebenitsch and C. Owen, "Individual variation in susceptibility to cybersickness," in *Proceedings of the 27th annual ACM symposium on User interface software and technology*, 2014, pp. 309–317.
- [39] T. C. Peck, L. E. Sockol, and S. M. Hancock, "Mind the gap: The underrepresentation of female participants and authors in virtual reality research," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 5, p. 1945–1954, 2020.
- [40] M. Q. Patton, "Qualitative research," *Encyclopedia of statistics in behavioral science*, 2005.
- [41] R. Li, M. van Almkerk, S. van Waveren, E. Carter, and I. Leite, "Comparing human-robot proxemics between virtual reality and the real world," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2019, pp. 431–439.
- [42] A. U. Batmaz, J. Maiero, E. Kruijff, B. E. Riecke, C. Neustaedter, and W. Stuerzlinger, "How automatic speed control based on distance affects user behaviours in telepresence robot navigation within dense conference-like environments," *Plos one*, vol. 15, no. 11, p. e0242078, 2020.
- [43] N. Dužmańska, P. Strojny, and A. Strojny, "Can simulator sickness be avoided? a review on temporal aspects of simulator sickness," *Frontiers in psychology*, vol. 9, p. 2132, 2018.

- [44] P. Kourtesis, S. Collina, L. A. A. Dumas, and S. E. MacPherson, "Validation of the virtual reality neuroscience questionnaire: maximum duration of immersive virtual reality sessions without the presence of pertinent adverse symptomatology," *Frontiers in human neuroscience*, vol. 13, p. 417, 2019.
- [45] D. Saredakis, A. Szpak, B. Birkhead, H. A. Keage, A. Rizzo, and T. Loetscher, "Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis," *Frontiers in Human Neuroscience*, vol. 14, 2020.
- [46] M. Slater, M. Usoh, and A. Steed, "Depth of presence in virtual environments," *Presence: Teleoperators & Virtual Environments*, vol. 3, no. 2, pp. 130–144, 1994.
- [47] M. Usoh, E. Catena, S. Arman, and M. Slater, "Using Presence Questionnaires in Reality," *Presence: Teleoperators and Virtual Environments*, vol. 9, no. 5, pp. 497–503, 10 2000. [Online]. Available: <https://doi.org/10.1162/105474600566989>
- [48] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3549–3557, Dec. 2009. [Online]. Available: <https://royalsocietypublishing.org/doi/10.1098/rstb.2009.0138>
- [49] J. Gugenheimer, D. Wolf, G. Haas, S. Krebs, and E. Rukzio, "Swivrchair: A motorized swivel chair to nudge users' orientation for 360 degree storytelling in virtual reality," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 1996–2000.
- [50] N. Matsuda, B. Wheelwright, J. Hegland, and D. Lanman, "Reverse pass-through vr," in *Special Interest Group on Computer Graphics and Interactive Techniques Conference Emerging Technologies*, 2021, pp. 1–4.