BRILL *Multisensory Research (2020)* DOI:10.1163/22134808-20201445 brill.com/msr

How Much of What We Learn in Virtual Reality Transfers to Real-World Navigation?

Lukas Hejtmanek [1,2], **Michael Starrett** [2,3,4], **Emilio Ferrer** [3] **and Arne D. Ekstrom** [2,3,4,*]

[1] Third Faculty of Medicine, Charles University, Ruská 87, Prague 10, 100 00, Czech Republic
[2] Center for Neuroscience, University of California, Davis, 1 Shields Ave, Davis, CA 95618, USA
[3] Department of Psychology, University of California, Davis, 1 Shields Ave, Davis, CA 95618, USA
[4] Department of Psychology, University of Arizona, 1503 E. University Blvd., Tucson, AZ 85719, USA

Received 15 February 2019; accepted 18 December 2019

Abstract

Past studies suggest that learning a spatial environment by navigating on a desktop computer can lead to significant acquisition of spatial knowledge, although typically less than navigating in the real world. Exactly how this might differ when learning in immersive virtual interfaces that offer a rich set of multisensory cues remains to be fully explored. In this study, participants learned a campus building environment by navigating (1) the real-world version, (2) an immersive version involving an omnidirectional treadmill and head-mounted display, or (3) a version navigated on a desktop computer with a mouse and a keyboard. Participants first navigated the building in one of the three different interfaces and, afterward, navigated the real-world building to assess information transfer. To determine how well they learned the spatial layout, we measured path length, visitation errors, and pointing errors. Both virtual conditions resulted in significant learning and transfer to the real world, suggesting their efficacy in mimicking some aspects of real-world navigation. Overall, real-world navigation outperformed both immersive and desktop navigation, effects particularly pronounced early in learning. This was also suggested in a second experiment involving transfer from the real world to immersive virtual reality (VR). Analysis of effect sizes of going from virtual conditions to the real world suggested a slight advantage for immersive VR compared to desktop in terms of transfer, although at the cost of increased likelihood of dropout. Our findings suggest that virtual navigation results in significant learning, regardless of the interface, with immersive VR providing some advantage when transferring to the real world.

* To whom correspondence should be addressed. E-mail: adekstrom@email.arizona.edu

 DOI:10.1163/22134808-20201445

Keywords
Navigation, virtual reality, spatial cognition, proprioception, transfer

1. Introduction

Many studies on the cognitive and neural basis of human spatial navigation involve virtual environments (VEs) rendered on desktop computers. A limitation of these studies is that they provide a limited approximation of the wealth of multisensory cues available during real-world navigation. Specifically, free ambulation involves critical body-based cues not represented during virtual reality (VR) navigation with a joystick on a desktop computer: cues derived from body turns and movements that displace fluid in our vestibular system (termed here 'vestibular cues') and also affect sensory receptors and proprioceptors in our body, particularly our legs (termed here 'proprioceptive cues') (Loomis and Beall, 1998; Starrett and Ekstrom, 2018). Vestibular input, though, is critical to our representation of bearing and acceleration, with lesions to the vestibular nuclei in both rats and humans significantly impairing navigation (Brandt *et al.*, 2005; Russell *et al.*, 2003; Taube *et al.*, 2013; Valerio and Taube, 2012). Similarly, although less well researched, proprioceptive cues associated with moving our legs and feet are important for estimating our velocity and turning angle (Loomis and Beall, 1998) and successful navigation more generally (Chrastil and Warren, 2013; Gallistel, 1990; Matthis *et al.*, 2018). In particular, the somatosensory responses of the feet are critical for estimating gait and other important aspects of movement in space, such as orientation (Lackner and DiZio, 2005; Matthis *et al.*, 2018; Visell *et al.*, 2011). Others have argued, however, that body-based cues may not all be necessary for normal expression of some forms of spatial representations, such as topological graph knowledge (Chrastil and Warren, 2015). At present, the degree to which body-based cues are necessary for spatial learning remains unresolved.

Immersive VR tools, such as head-mounted displays (HMDs) and omnidirectional treadmills, allow a novel opportunity to study immersive navigation in the lab and more completely model body-based input in a controlled setting. Specifically, omnidirectional treadmills (see Fig. 1) allow participants to freely ambulate while wearing an HMD with the full spectrum of rotations that can render any number of different VEs while presenting them with a wide field of view (about 100 degrees in the HTC Vive HMD (HTC Corp., Taoyuan, Taiwan) in comparison to 60–70 degrees on a desktop monitor). With these two important components, we can study both enriched visual input (*via* the HMD) and the effect of enriched body-based cues on navigation (*via* the omnidirectional treadmill). Specifically, the omnidirectional treadmill

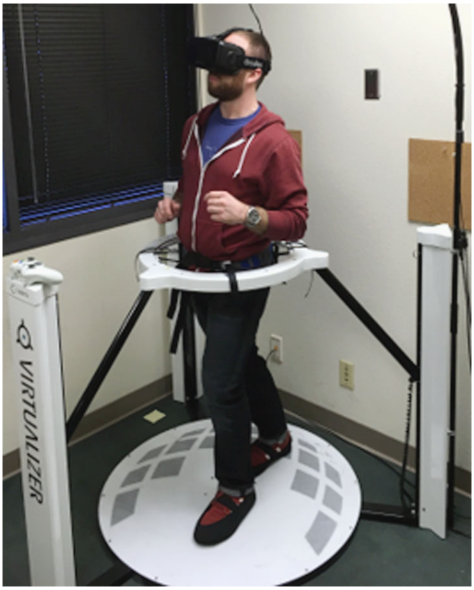

Figure 1. Immersive VR treadmill setup. Participant wearing an HTC Vive headset while being strapped into the omnidirectional treadmill. Participants can rotate 360 degrees and are allowed independent body and head rotation. Movement is achieved by sliding one's feet across the treadmill base.

provides both body and turn input, proprioceptive input from the legs, and somatosensory feedback from the feet as participants walk on the surface of the treadmill. Previous studies investigated navigation by providing participants with a full range of proprioceptive cues using an HMD and ambulation in a room (Ruddle and Lessels, 2006). Although omnidirectional treadmills only approximate the real-world experience (for example, walking on the omnidirectional treadmill does not engage the full range of somatosensory receptors and muscles compared to stepping in the real world), they have the advantage of allowing participants to explore environments larger than a VR/AR-enabled room would offer. Therefore, an important question regards to what extent information learned in such immersive VR contexts compares to real-world navigation.

One way to assess how well information acquired during virtual navigation applies to the real world involves 'transfer'. This typically involves subjects first learning environments in a virtual setting and then determining how effectively participants can apply acquired spatial knowledge to the real world (Montello *et al.*, 2004; Richardson *et al.*, 1999; Waller *et al.*, 1998; Witmer *et al.*, 1996). Studies investigating transfer of spatial knowledge acquired in VR to the real world also suggest advantages for conditions involving some amount of body-based input. One study by Waller *et al.* compared continuous exposure to an environment on a desktop VE, navigating the same environment with an HMD with full head control and a joystick, and navigating the same real-world environment. HMD exploration with full head turning led to greater transfer than desktop exploration, although real-world navigation led to the highest transfer (best spatial learning). These findings are similar to some reported in other past studies (Waller *et al.*, 1998; Witmer *et al.*, 1996; see also Richardson *et al.*, 1999), suggesting the importance of body-based cues to navigation and transfer of information to real-world environments. One possibility, though, is that because these studies did not involve walking in the VR conditions, the limited proprioceptive cues led to overall poor encoding of head direction information due to a mismatch with the joystick movements. Thus, we might expect that the presence of more enriched body-based cues could lead to better transfer to the real world, an issue we will address in this study.

In support of the importance of richer body-based cues to navigation, Grant et al. compared navigation with an HMD, head turns, and walking in place by shuffling the feet compared to movement with a joystick (Grant and Magee, 1998). While the HMD/walking condition did result in improvements in taking shortcuts compared to desktop VE, both conditions resulted in worse direction estimates compared to real-world navigation. One limitation of the Grant and Magee study, however, is that the seated shuffling movement of the feet they employed should read out to the proprioceptive system as a net-zero forward translation. In contrast, on an omnidirectional treadmill, standing and moving the feet in a walking motion should provide a sense of net-positive forward translation. Thus, while studies of transfer from VR to real-world environments suggest advantages to some vestibular input rendered by an HMD, the absence and mismatch, in some cases, with other body-based cues (i.e., walking-based input from sensory and muscle receptors) may be a possible reason for incomplete transfer. Alternatively, it is also possible that visual rendering in VR always limits transfer (Thompson *et al.*, 2004), and thus it is possible that, even with richer body-based cues, visual rendering always limits transfer, to some extent.

In this study, we contrast two different forms of learning in VR and how they transfer to real-world navigation. As a control comparison and means of determining the theoretical ‘maximum’ transfer, we included continuously navigating in the real world for the entire experiment. To avoid issues with navigating in outdoor environments (trip hazards and environmental complications, like variable levels of lighting and rain), all testing occurred in a campus building, the UC Davis Center for Neuroscience (Figs 2, 3). One-third of our participants navigated through the building on desktop, the second third in immersive VR (i.e., HMD + omnidirectional treadmill), and the final third of

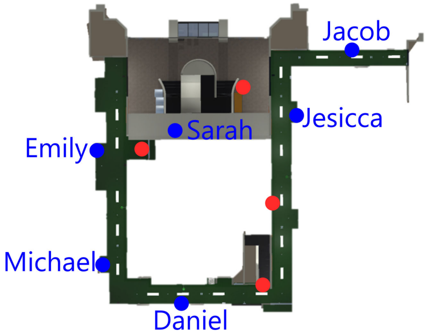

Figure 2. Illustration of the building layout with six offices (dots with names in blue) and four pointing locations.

Figure 3. Comparison of the visual and geometrical fidelity for the 3D visualisation. Center for Neuroscience entrance hall is depicted on the left and the modelled environment on the right.

participants navigated the real-world Center for Neuroscience. After completing three blocks of navigation (by finding specific targets in the building), all participants navigated the Center for Neuroscience in the real world. By tracking their position in the real-world building using Bluetooth trackers beacons (Estimote Inc., San Francisco, CA, USA), we were able to determine the deviation of their path from the ideal path they should walk, as well as any errors in visiting the wrong target.

Our main hypothesis, which we term the body-based enrichment hypothesis, suggests that providing additional body-based cues with an omnidirectional treadmill and HMD should at least partially mimic real-world navigation. In this case, we would predict that transfer should be higher from immersive VR to the real world than desktop VR. Immersive VR, according to the body-based enrichment hypothesis, should result in comparable, or maybe less transfer compared to the real world, depending on the degree to which participants experience the immersive experience as fully mimicking the real world. To briefly preface our results, we found that learning in the real-world setting led to the highest transfer overall, with some evidence for immersive VR leading to greater transfer than desktop navigation.

2. Methods

2.1. Design Overview

To investigate the effects of learning modality on spatial knowledge acquisition, we devised an experiment with three different conditions involving active exploration and ambulation. Participants learned positions of six office doors (out of a total 32) within the Center for Neuroscience at UC Davis. The reason we had participants learn the offices was to avoid them having to remember objects they could easily see. Participants experienced the same building either in the real world (i.e., actually navigating the Center for Neuroscience), in immersive body-based VR with an omnidirectional treadmill (i.e., navigating a virtual version of the Center for Neuroscience with an HMD on an omnidirectional treadmill), or on a desktop computer (i.e., navigating the same virtual version but on a desktop computer with a mouse and a keyboard). After three blocks of learning, all participants continued the task in the real world. Our dependent measures involved (1) normalized walked path distance, which we assayed in the (a) real world using Estimote Bluetooth trackers and (b) in VR by recording position (termed 'path accuracy'); (2) visitation errors, which were quantified based on visiting the wrong doors; and (3) unsigned pointing error.

2.2. Walking Task

At the beginning of each walking trial, participants were given the name of a person whose office they had to find. Instructions were provided by the experimenter in the real-world condition or displayed in the user interface in the treadmill VR and desktop conditions. Participants selected the door by walking directly into it in virtual conditions and standing close and pointing to it in the real-world condition. Participants were not told which office belonged to whom, and therefore they had to walk around the building and keep checking doors until they arrived at the correct one. Once they arrived or pointed at the correct door participants were provided with a new name and a new trial started. Each office was visited three times in each phase, six times over the entire experiment. The order in which offices occurred was randomized in a way such that the order was balanced across all conditions.

2.3. Pointing Task

After finishing three blocks of finding doors (the learning phase), the participant was either moved (VR and desktop conditions) or walked to two different viewpoints within the building and asked to point to each office one by one.

This was repeated once more after the transfer phase, for a total of four pointing blocks (24 pointing in total).

2.4. Building a Realistic 3D Model of the Center for Neuroscience

We built the virtual Center for Neuroscience to scale for rendering in both immersive VR and on the desktop VE. An example of a viewpoint in the virtual rendition and the real world can be seen in Fig. 3. We set the walking speed in the desktop so that walking the length of a corridor in the real world would take the same time as walking the corridor in the desktop version. In the immersive VR condition, we set the speed so that walking down the corridor would take approximately the same number of steps in real world as it did on the treadmill. The experiment was built in Unity3D (Unity Technologies, San Francisco, CA, USA). The building can be experienced online at https://hejtmy.github.io/CFNS-task/. For the virtual conditions in the immersive VR condition, we used the HTC Vive HMD to render the building and participants moved using the Cyberith Virtualizer omnidirectional treadmill (Cyberith GmbH, Herzogenburg, Austria) to provide body-based cues (Fig. 1). Participants had a full range of rotational motion available to them. The desktop condition was presented on a 21″ monitor (resolution 1920×1080) and participants controlled it with a keyboard and a mouse. All participants were informed about possible side effects of the VR and in case of uncomfortable or prolonged dizziness, such participants were dismissed from the study and their data removed. Participants in the immersive VR condition in the learning phase were significantly more likely to drop out due to the cybersickness, fatigue, or other factors related to discomfort. We observed this both during the first [$\chi^2(1, 82) = 9.91$, $p = 0.002$] and the second experiment [$\chi^2(1, 70) = 14.72$, $p < 0.001$]. Detailed dropout rates for all conditions can be found in Tables 1 and 2.

Table 1.
Cyber sickness dropout table for Experiment 1 (all retests in the real world): Finished/total (percent finished). One participant in each group was removed during analyses due to recording failure, but they finished the experiment without issues. Participants in the immersive VR condition in the first phase were significantly more likely to drop out [$\chi^2(1, 82) = 9.91$, $p = 0.002$]

Learning condition	Male	Female
Real world	4/5 (80%)	19/19 (100%)
Desktop	11/11 (100%)	9/11 (82%)
Immersive VR	9/13 (69%)	12/23 (52%)

Table 2.
Cyber sickness dropout table for Experiment 2 (all transfer phase in the VR): Finished/total (percent finished). One participant from the immersive VR group and two from the real-world group were removed during analyses due to recording failure, but they finished the experiment without issues and one participant in the immersive VR condition who did not finish did not disclose their gender. Participants in the immersive VR condition in the first phase were significantly more likely to drop out [$\chi^2(1, 70) = 14.721$, $p < 0.001$]

Learning condition	Male	Female
Real world	8/9 (89%)	14/16 (88%)
Immersive VR	8/17 (47%)	13/27 (48%)

2.5. Real-World Tracking

We tracked participants' position within the building using Bluetooth and a custom-made app running on the iPhone 6. Participants were given the phone to hold directly in front of them to track both position and rotations of their body. As participants walked around the building, the app tracked the strength of the Bluetooth signal towards the closest beacons and triangulated position within given building constraints. Participants were accompanied by an experimenter who provided instructions and feedback during door selection and noted down errors and special events. For the ambulatory paths travelled in the real world as well as on the VR treadmill, trajectories were smoothed using a median value in a 5-s moving window to remove slight jitter introduced by fluctuating strength of the Bluetooth signal in the real world or the shuffling of the feet on the treadmill.

2.6. Dependent Measures

We calculated participants' walked distance in the building, time spent in each task, and the number of incorrectly visited doors during walking trials. For the pointing tasks, we calculated unsigned pointing error. The three conditions were not directly comparable in their measured distance and time, with the VR trials taking a little bit longer in both time and distance than the real-world condition. The increased time was due to the novelty of the movement control that persevered even after the training session. In contrast, the larger absolute distance was due to differences in tracking precision of the VR and desktop compared to the real-world setup. We also observed feet shuffling and occasional small stumbles on the treadmill that added to the absolute walked distance for immersive VR in comparison to the desktop condition.

We therefore min-normalized the distance travelled and the time. We used the shortest path for each task (pair of offices, e.g., Jacob to Sarah) that any participant demonstrated in a single environment (e.g., real world) and then divided other participants performance in the same task (Jacob to Sarah) and in the same environment (real world) by this number. Therefore, we had three baselines (desktop, immersive VR and real world) for each task (pair of offices). For example, a min-normalized distance value of 2.00 would mean the participant travelled double the distance than what was the best performance in each task (pair of offices) and in a certain modality (desktop, immersive VR or real world). Unless stated otherwise, we calculated all statistics on these normalized measures. We removed trials where distance or time measurements were more than three standard deviations away from the mean (a total of 1.80 percent of trials).

2.7. Experimental Procedure

Participants walked to six different office doors in the Center for Neuroscience building at UC Davis, in real or virtual versions. Positions of all the office doors can be seen in Fig. 2. The task itself consisted of a set of 36 walking trials and 24 pointing trials, separated into two equal phases, the learning phase and the transfer phase. Each phase had 18 walking trials consisting of three different blocks (each office visited once per block, three times per phase). After each phase, participants did a short pointing session of 12 trials (pointing to each of the six offices from two viewpoints). In the transfer phase, the immersive VR and desktop group switched their environment modality to the real world, whereas the real-world group remained in the same environment. Each participant, regardless of experimental condition, received a short training session on the VR treadmill before the learning phase.

Because our initial analyses observations and analyses suggested that participants might learn more slowly in VR, we devised a second experiment to better understand this issue. One group learned the UC Center for Neuroscience building in immersive VR while the second group learned in the real world. They both then transferred to immersive VR. Although this experiment does not directly address our main hypothesis (transfer from immersive/desktop VR to reality), this allowed us to assess if the immersive VR can achieve the same performance given additional time. It also allowed us to look at the conceptual ‘inverse’ process of transfer.

Schema of the entire procedure can be seen in Fig. 4.

2.8. Participants (Experiment 1)

A total of 82 undergraduate students at the University of California, Davis participated in the study in exchange for a course credit. Each participant was randomly assigned a condition and a randomized set of goals before arrival. Eighteen participants did not finish due to motion sickness and three were removed due to technical problems with the real-world tracking system. This resulted in a final sample size of 61 participants (37 female) (age $M = 20.42$, $\mathrm{SD} = 2.16$) used for all reported analyses. The participant completion rates for

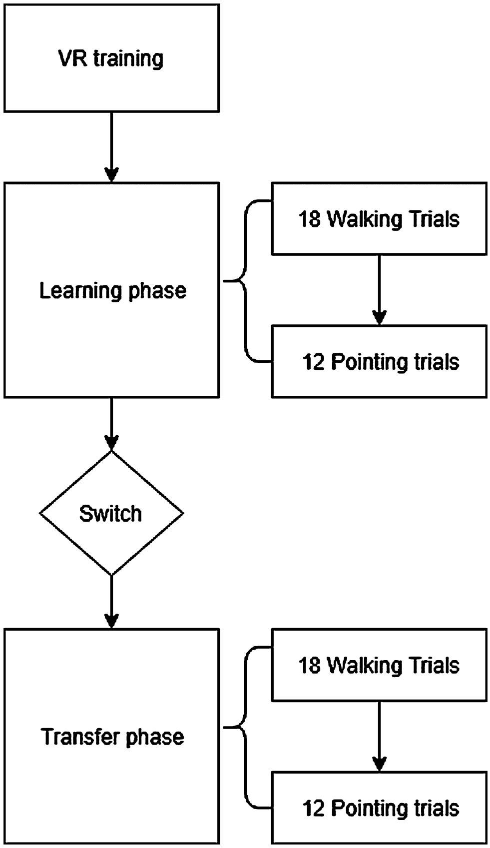

Figure 4. Visualisation of the experimental procedure. Session consisted of training, learning and transfer phases. Both learning and transfer phases included 18 walking trial and 12 pointing trials.

Experiment 1 are in Table 1. All procedures were approved by the UC Davis Institutional Review Board (IRB).

2.9. Participants (Experiment 2)

A total of 70 undergraduate students at UC Davis participated in the second experiment. Twenty-nine participants did not finish due to motion sickness or other issues with the VR interfaces. Analyses were then conducted on a final set of 41 (14 female) participants (age $M = 19.96$, SD $= 1.55$).

2.10. Analyses

We analysed the data in R (R Core Team, R Foundation for Statistical Computing, Vienna, Austria), with the help of the *ez* package for ANOVAs and the ggplot package for plots. Based on our *a-priori* hypothesis that real world $\geqq$

immersive > desktop VR, we performed t-tests uncorrected for multiple comparisons. This allowed us to explore whether effects were consistent with predictions. To determine learning rates (slopes) over the experiment, we performed mixed-model analyses of the entire experiment in MPlus (https://www.statmodel.com/company.shtml).

3. Results

3.1. Experiment 1: Quantifying Learning Rates Prior to Transfer (Blocks 1 to 3)

We first considered blocks 1–3, when all participants navigated in separate modalities (i.e., desktop, immersive, or real world). Considering the very first block of exposure to the environment, we found that participants in the different learning modalities did not differ in their performance either in normalized distance [$F(2, 323) = 1.99$, mean squared error (MSE) $= 16.13$, $p = 0.139$, $\hat{\eta}^2_G = 0.012$] nor the number of visitation errors [$F(2, 363) = 0.10$, MSE $=$ 64.52, $p = 0.903$, $\hat{\eta}^2_G = 0.001$]. Some caution is needed, however, with these comparisons because participant variability was likely high on the first exposure to the environment. The conditions also did not differ in the second block for normalized distance [$F(2, 317) = 0.47$, MSE $= 4.61$, $p = 0.627$, $\hat{\eta}^2_G = 0.003$] although we saw a significantly lower number of visitation errors for real-world group ($M = 4.24$) compared to immersive-VR ($M =$ 7.54) [$t(249.98) = 4.01$, $p < 0.001$, $d = 0.504$] and desktop ($M = 8.83$) [$t(241.41) = 5.39$, $p < 0.001$, $d = 0.687$].

We then compared the last block before transfer (block 3) with independent sample t-tests. We found a significant difference in visitation errors between real-world and desktop learning [$t(198.7) = 6.84$, $p < 0.001$, $d = 0.896$] and real-world and immersive VR learning [$t(179.9) = 7.40$, $p <$ 0.001, $d = 0.958$], but no difference between desktop and immersive learning [$t(223.1) = 1.44$, $p = 0.150$, $d = 0.188$]. The comparison was the same for normalized distance: we found a significant difference in distances travelled in block 3 between real-world and desktop conditions [$t(161.80) = 1.99$, $p =$ 0.048, $d = 0.259$] and real-world and immersive learning [$t(140.58) = 3.09$, $p = 0.002$, $d = 0.389$], but no difference between desktop and immersive learning [$t(203.90) = 1.42$, $p = 0.156$, $d = 0.187$]. These findings indicate that prior to transfer (blocks 1–3), both immersive VR and desktop resulted in less spatial knowledge than real-world navigation, particularly for visitation errors. These findings suggested that participants in both VR conditions ended the learning phase performing slightly but significantly worse than real-world participants, particularly for visitation errors.

Our findings for pointing errors (which were only collected on blocks 3 and 6) echoed the same basic pattern described above for normalized path and visitation errors. We again observed significant differences between conditions in pointing performance at the end of the first phase [$F(2, 729) = 17.78$, MSE $= 1412.53$, $p < 0.001$, $\hat{\eta}_G^2 = 0.047$] (Fig. 5c). Specifically, participants in the real-world learning group ($M = 21.93$, SD $= 26.34$) performed significantly better compared to both the immersive VR group ($M = 35.23$, SD $=$ 37.48) [$t(424.16) = 4.57$, $p < 0.001$, $d = -0.414$] and the desktop group ($M = 41.62$, SD $= 47.47$) [$t(342.85) = 5.57$, $p < 0.001$, $d = 0.523$]. Numerically, this trend followed real world > immersive > desktop in terms of errors, although the differences between desktop and immersive VR were not statistically significant ($p = 0.108$). Overall, these findings suggested that participants acquired information most readily in the real world, although significant degrees of learning occurred in all three conditions, with the immersive condition showing at least numerically better pointing error than desktop.

3.2. Transfer Effects (Block 3 to Block 4)

The critical area of interest in our study related to how participants transferred information from VR modalities to the real world. We addressed this issue by directly comparing normalized distance and visitation errors on the last block before the transfer phase (block 3) with the first block of the transfer phase (block 4). Visitation errors and normalized distance were each entered into a 3 (Learning Condition: real world, desktop, immersive VR) $\times$ 2 (Block: block 3, block 4) mixed-measures ANOVA. For visitation errors, there was a main effect for block [$F(1, 58) = 42.67$, MSE $= 2.19$, $p < 0.001$, $\hat{\eta}_G^2 = 0.112$], a main effect of condition [$F(2, 58) = 13.66$, MSE $= 10.61$, $p < 0.001$, $\hat{\eta}_G^2 =$ 0.281] and a condition by block interaction effect [$F(2, 58) = 9.22$, MSE $=$ 2.19, $p < 0.001$, $\hat{\eta}_G^2 = 0.052$]. For normalized distance, there was a main effect of condition [$F(2, 58) = 6.65$, MSE $= 0.90$, $p = 0.003$, $\hat{\eta}_G^2 = 0.110$], but no effect of block [$F(1, 58) = 2.61$, MSE $= 0.78$, $p = 0.111$, $\hat{\eta}_G^2 = 0.020$] and the interaction also failed to reach significance [$= F(2, 58) = 1.89$, MSE $=$ 0.78, $p = 0.161$, $\hat{\eta}_G^2 = 0.029$]. These findings suggest that only visitation errors showed a significant improvement during transfer, unhindered by the change of modality, although the transfer modality (desktop *vs* immersive *vs* real world) impacted the degree of transfer.

We then ran pairwise t-tests to compare the improvement in normalized distance and visitation errors change from block 3 to block 4. For visitation errors, there were significant changes for all conditions, with effect sizes higher (in other words, more transfer) for going from the immersive VR to the real world ($d = 1.33$) compared to desktop to the real world ($d = 0.56$) (see Fig. 6a and Table 3). Here, a larger effect size means a greater reduction from block 3 to

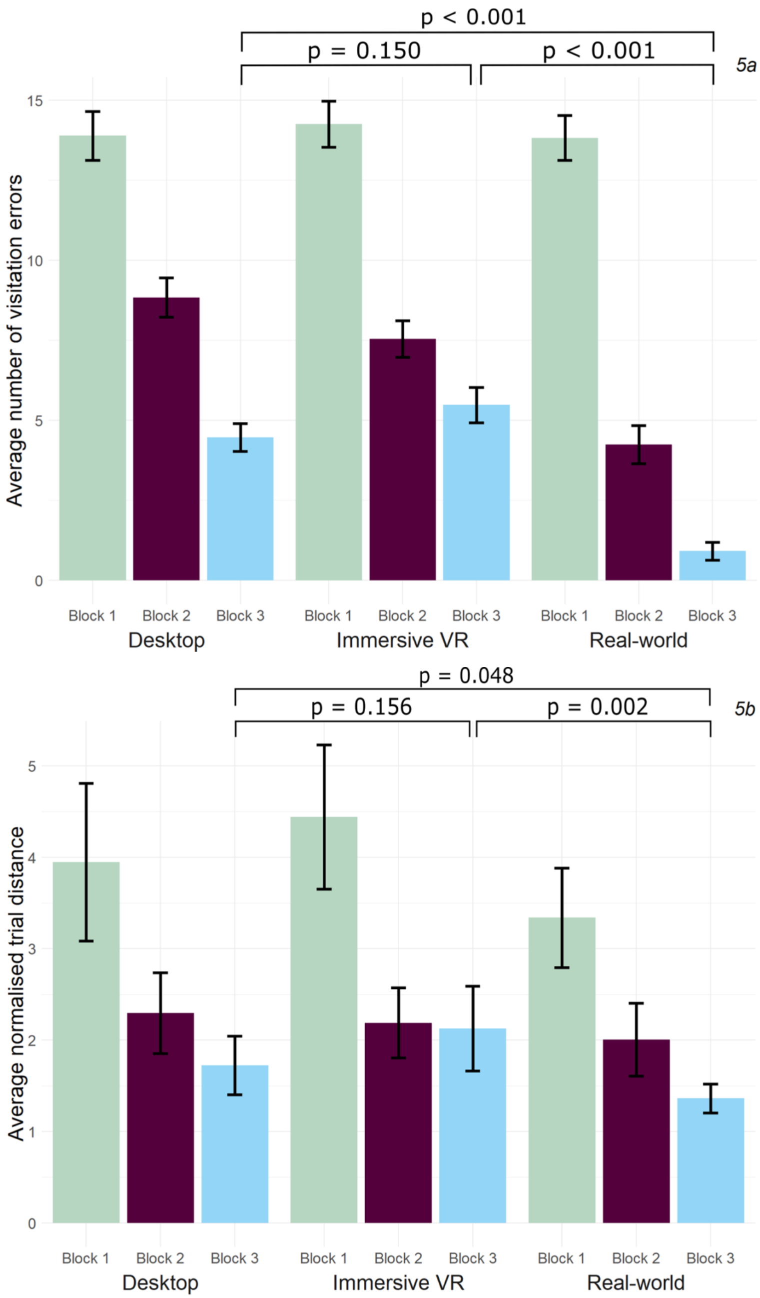

Figure 5. (a) Average number of visitation errors per block in the learning phase (blocks 1 to 3) for Experiment 1, split by learning condition (SEM error bars). (b) Average normalised distance per block in the learning phase (blocks 1 to 3) for Experiment 1, split by learning condition (SEM error bars). (c) Average absolute pointing error at the end of the learning phase for Experiment 1, split by learning condition (SEM error bars).

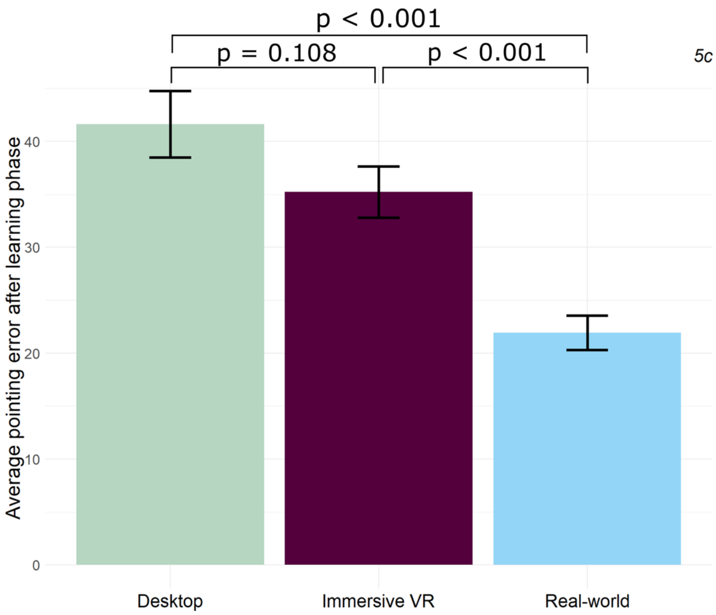

Figure 5. (Continued.)

4 and thus more transfer, suggesting that the immersive conditions resulted in greater transfer. Note that while the effect sizes were lower for real world to real world ($d = 0.55$); this was likely because participants in this condition were already significantly better at real-world navigation by block 3 (0.91 errors on average compared to 5.47 in VR and 4.46 on the desktop), and thus had less to learn. In contrast, for normalized distance, there were no significant changes from block 3 to block 4 (see Fig. 6b and Table 4). These findings suggest a slight advantage for transfer from immersive VR compared to the desktop.

Because an important goal for similar projects is to study participant's performance after being trained under different conditions of immersion, we compared performance in the first block of transfer (block 4). Because all participants navigated the real building during block 4, this also allowed us to control for potential differences between learning conditions and remove potential confounds for block 3 to block 4 comparisons. During block 4, we observed significant differences in visitation errors when comparing real world to the desktop [$t(130.84) = 5.21$, $p < 0.001$, $d = 0.709$] and immersive VR groups [$t(141.81) = 3.74$, $p < 0.001$, $d = 0.491$]. In this case, greater effect sizes between conditions mean a greater difference, i.e., a larger discrepancy between real world and desktop than real world and immersive VR. This suggests again that the immersive condition was closer in terms of transfer,

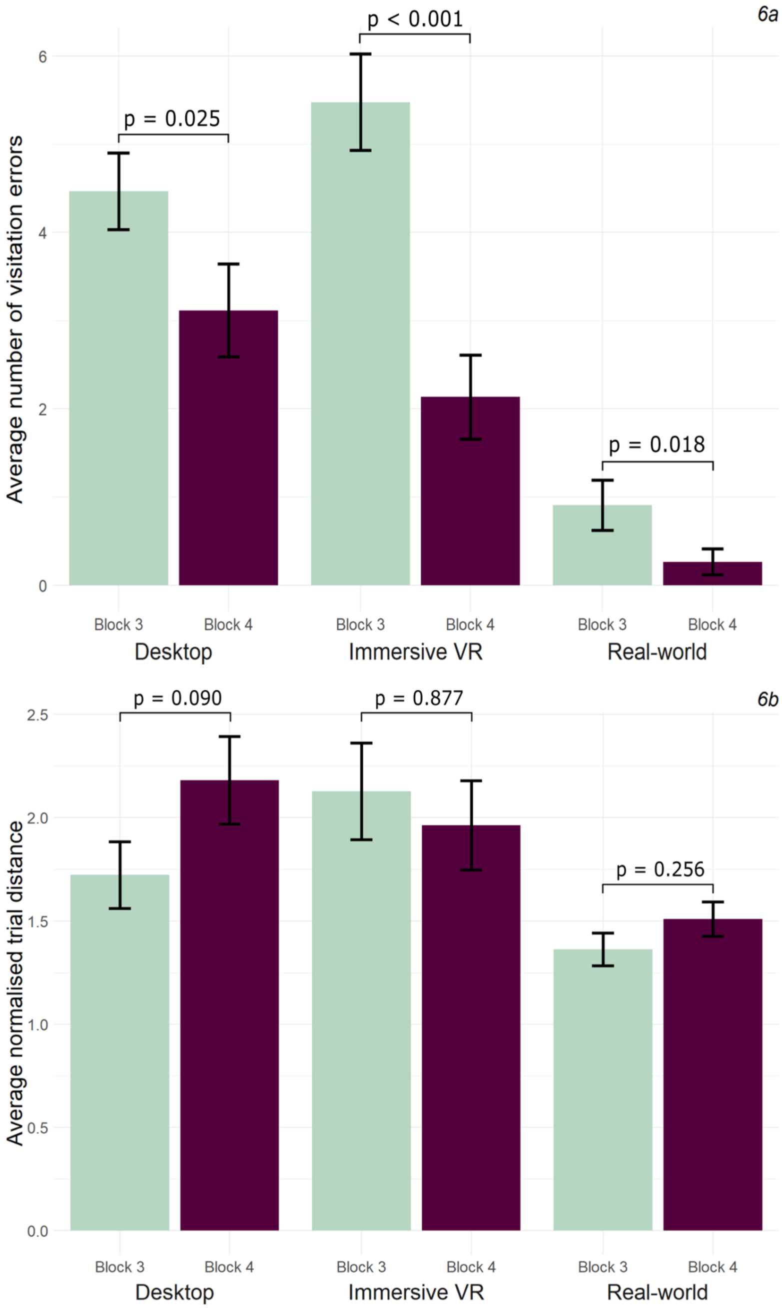

Figure 6. (a) Average number of visitation errors in blocks 3 and 4 (last learning block to first transfer block), split by learning condition (SEM error bars). (b) Average normalised distance in blocks 3 and 4 (last learning block to first transfer block), split by learning condition (SEM error bars).

Table 3.
Visitation error improvements across the three learning conditions from block 3 to block 4 — all conditions transfer to the real world in block 4

Learning condition	M	SD	M_{diff}	df	t	p	d	95% CI
Real world								
Block 3	0.91	3.27	0.64	21	2.56	0.018	0.55	0.12, 1.17
Block 4	0.27	1.70						
Desktop								
Block 3	4.46	4.65	1.35	18	2.44	0.025	0.56	0.19, 2.52
Block 4	3.11	5.62						
Immersive VR								
Block 3	5.47	6.00	3.34	19	5.97	<0.001	1.33	2.17, 4.51
Block 4	2.13	5.22						

Table 4.
Normalized distance improvements across the three learning conditions from block 3 to block 4 — all conditions transfer to the real world in block 4

Learning condition	M	SD	M_{diff}	df	t	p	d	95% CI
Real world								
Block 3	1.36	0.75	−0.15	21	−1.17	0.256	−0.25	−0.41, 0.12
Block 4	1.51	0.87						
Desktop								
Block 3	1.72	1.73	−0.70	18	−1.79	0.090	−0.41	−1.53, 0.12
Block 4	2.18	2.02						
Immersive VR								
Block 3	2.13	2.52	0.04	19	0.16	0.877	0.04	−0.55, 0.63
Block 4	1.96	2.09						

based on effect sizes, to the real world than desktop. A similar relationship was observed in normalized distance, with participants' learning in the real world performing better than on desktop [$t(118.02) = 2.95$, $p = 0.004$, $d = 0.446$] and immersive VR, although this comparison failed to reach significance [$t(119.40) = 1.96$, $p = 0.052$, $d = -0.292$]. Note, however, that immersive and desktop VR did not differ for either dependent measure [visitation errors: $t(228.45) = 1.38$, $p = 0.168$, $d = 0.181$] or normalized distance [$t(181.98) = 0.72$, $p = 0.472$, $d = 0.106$]. Given that the effect sizes suggested larger differences for desktop VR *vs* real world compared to immersive VR *vs* real world, our findings again point to a slight advantage for immersive VR in terms of transfer.

3.3. Final Performance (Block 6)

To understand how the learning modality affected the last block performance, we compared dependent measures on the 6th block, after each group had finished three blocks of navigation in the real world. Because we also collected pointing accuracy on the final block, we also included this measure. In the final testing block (block 6), we found no differences among the groups in either visitation errors [$F(2, 363) = 1.67$, MSE $= 3.33$, $p = 0.191$, $\hat{\eta}^2_G = 0.009$] nor pointing performance [$F(2, 726) = 1.56$, MSE $= 982.05$, $p = 0.210$, $\hat{\eta}^2_G = 0.004$]. These findings suggested that by the 6th repetition, all groups had reached comparable levels on visitation errors and pointing accuracy. We did find, however, a small but significant difference in normalized distance when comparing all three conditions in an ANOVA [$F(2, 255) = 3.21$, MSE $= 0.10$, $p = 0.042$, $\hat{\eta}^2_G = 0.025$]. Specifically, by the last block, we found that the real-world group ($M = 1.25$, SD $= 0.19$) performed better than immersive VR group ($M = 1.37$, SD $= 0.38$) [$t(102.13) = 2.55$, $p = 0.012$, $d = -0.426$], but not better than desktop group ($M = 1.28$, SD $= 0.37$) [$t(131.62) = 0.90$, $p = 0.369$, $d = 0.135$]. For visitation and pointing errors, these findings support the idea that all participants reached the same level of knowledge by the last block. For normalized distance, our findings suggest that real-world learning continued to confer an advantage.

3.4. Modelling Changes Over All Six Learning Blocks

Although we performed t-tests earlier based on our *a-priori* hypotheses, as explained in the introduction, we thought it also important to look at whether subjects learned over the entire experiment. We employed a mixed-effects linear model over all six learning blocks, which involved looking at changes in normalized distance and visitation errors as separate dependent measures for Experiment 1. For visitation errors and normalized path, all three conditions showed significant negative slopes over six blocks of learning, indicating that all three conditions resulted in improvements in learning (desktop → real world, immersive VR → real world, and real world → real world; Supplementary Table S1). For visitation errors, we observed an interaction effect [difference in slopes: $t(3565) = 5.38$, $p < 0.0001$] while for normalized path, there was no significant difference in slopes. The interaction effect suggested that the conditions were learned at different rates across the three conditions, which were explored earlier with t-tests.

3.5. Experiment 2: Comparing Transfer of Real World to Immersive VR

Findings from Experiment 1 suggested that the VR interfaces, particularly the treadmill, might be difficult for participants to learn as quickly as the real world. Experiment 2 assessed the ‘opposite’ transfer (both conditions transferred to the immersive VR) and we therefore investigated differences between the two learning conditions (real world, immersive VR) in the last (6th) block using t-tests. The conditions did not differ in normalized distance [$t(152.12) = 0.88$, $p = 0.378$], visitation errors [$t(228.89) = 0.29$, $p = 0.771$, $d = 0.037$], nor pointing error [$t(485.14) = 0.17$, $p = 0.863$, $d = -0.016$]. These findings suggested that both immersive VR and real-world learning lead to the same eventual performance when transferring to immersive VR. Comparing both conditions immediately after transfer to the immersive VR (block 4), we found a trending relationship in average visitation errors [$t(240.09) = 1.90$, $p = 0.058$, $d = 0.242$], with the real-world learning group committing fewer errors ($M = 3.28$) than the immersive VR learning group ($M = 4.31$), but no difference for normalized distance [$t(241.53) = 0.46$, $p = 0.647$, $d = 0.059$].

We then compared all conditions from Experiment 1 (real world, immersive VR and desktop transferring to the real world) and Experiment 2 (real world and immersive VR transferring to the immersive VR). We found a significant effect of condition on final normalized distance [$F(4, 498) = 3.68$, $\mathrm{MSE} = 0.19$, $p = 0.006$, $\hat{\eta}^2_G = 0.029$], with Tukey post-hoc tests revealing a significant difference between the real world to immersive VR group *vs* immersive VR to real world ($p < 0.001$) (see Fig. 7a and Table 5). No other comparisons reached significance for normalized distance. For visitation errors, we found significant differences between conditions in the last block across experiments [$F(4, 607)$, $\mathrm{MSE} = 4.62$, $p < 0.001$, $\hat{\eta}^2_G = 0.181$]. Tukey's post-hoc tests suggested that the difference was due to the transfer phase modality, e.g., participants in Experiment 1 who learned the real-world condition in the transfer phase always performed better than participants in Experiment 2, who had immersive VR in the transfer phase (see Fig. 7b and Table 5).

Together, these results are consistent with our earlier findings suggesting that real-world navigation, particularly real world to real world navigation, conferred a 'transfer' advantage compared to all VR conditions. These findings also reinforce the idea that, despite capturing some aspects of real-world navigation, the treadmill interface failed to fully substitute for it. The findings, however, argue against the idea that our results are an artefact of difficulty with the treadmill interface, as participant clearly learned and transferred information from the real world to immersive VR and *vice versa*.

4. Discussion

In our experiment, we were able to directly compare acquisition of spatial knowledge in three different modalities: navigating the real world, navigating an immersive virtual interface involving walking on a treadmill and viewing

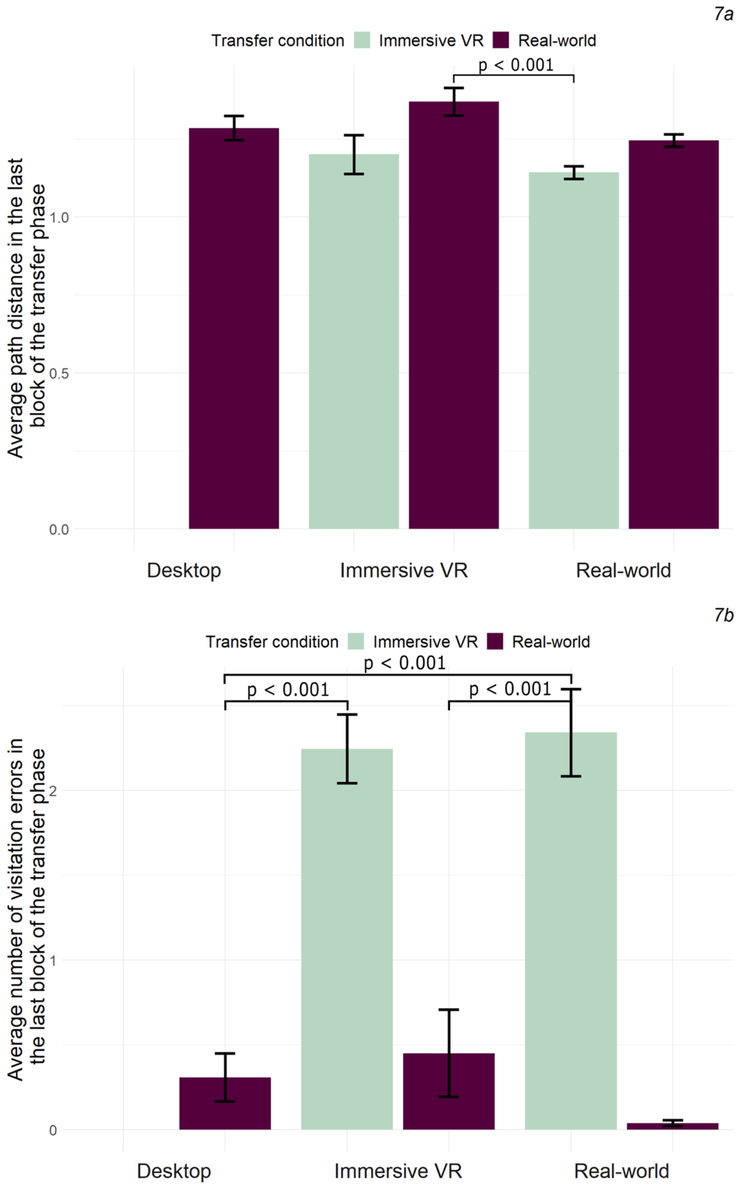

Figure 7. (a) Mean normalised distance in the last block (block 6) for each combination of learning phase (blocks 1–3, bottom labels) and transfer phase (blocks 4–6) conditions. (b) Average number of visitation errors made in the last block (block 6) for each combination of learning phase (blocks 1–3, bottom labels) and transfer phase (blocks 4–6) conditions.

Table 5.
Mean normalised distances and visitation errors in the last transfer block for both experiments. Standard deviation in parentheses

Learning condition	Transfer condition	Last block mean normalised distance	Last block mean visitation errors
Real world	Real world	1.25 (0.19)	0.04 (0.19)
Desktop	Real world	1.28 (0.37)	0.31 (1.51)
Immersive VR	Real world	1.37 (0.38)	0.45 (2.82)
Immersive VR	Immersive VR	1.20 (0.7)	2.25 (2.28)
Real world	Immersive VR	1.14 (0.22)	2.34 (2.82)

the environment on an HMD, and navigating the same environment on a desktop computer with a joystick. We investigated how navigating in these three different conditions 'transferred' spatial knowledge to the same real-world environment. To fully capture the transfer process, we employed two different dependent measures interspersed with each of the six different blocks, visitation errors (visiting the wrong door) and normalized path; we also employed one dependent measure every three blocks, pointing error. For the first three blocks of learning prior to transfer, our findings show an advantage for real-world navigation compared to both immersive VR and desktop navigation on all three dependent measures. These findings suggest that participants learned the most information during real-world navigation, although, in all cases, they also showed improvements on visitation errors and normalized distance over the first three blocks in the two virtual navigation conditions, suggesting they acquired spatial information effectively from these two modalities as well.

After completing the first three blocks of navigation in a specific modality, participants transferred to the real world. Particularly for visitation errors, we found differences as a function of modality. While both virtual conditions showed less transfer to the real world than prior real-world navigation, we observed a greater decrease in visitation errors (higher effect size) during transfer for the immersive VR compared to the desktop condition. We also found that immersive VR resulted in numerically lower pointing errors after three blocks of navigation than desktop VR (Fig. 5c), although the different was not statistically significant. These findings are consistent with the idea that immersive technologies, particularly those capturing rotational and somatosensory cues associated with walking, lead to some improvements in learning compared to desktop VE (Chrastil and Warren, 2013; Grant and Magee, 1998; Klatzky *et al.*, 1998; Ruddle and Lessels, 2006; Waller *et al.*, 2004). At the same time, we did not find differences in transfer for normalized distance. We measured normalized distance using positional information in virtual conditions and Bluetooth tracking using an iPhone in the real world. Comparing normalized distance from the second experiment involving navigating in the real world and transferring to immersive VR (or continuously navigating in immersive VR) similarly showed no differences in normalized distance, although continuous navigation in the real world did show an advantage over any navigation involving immersive VR.

Comparing the results from both experiments, we established that all learning conditions reached comparable levels of performance in the end. But we also observed that visitation errors in the last block were lower in the real world than in immersive VR. Also, participants learning in immersive VR and transferring to the real world demonstrated worse walking performance than participants transferring to immersive VR from the real world. Our findings therefore corroborate and elaborate on previous studies (Farrell *et al.*, 2003) which suggest that path and directional knowledge can be learned and transferred from VR, in many cases, quite effectively for building-sized environments, although real-world navigation still confers some advantages.

The effects we observed in terms of learning rates and transfer were more pronounced for visitation errors than normalized path. Why would we see differences in memory for specific doors for the different virtual conditions but less so, and not at all in some cases (transfer metric) for path error? Notably, the environment we tested was a medium sized campus building, the Center for Neuroscience. It is possible that the paths were relatively easier for subjects to master, and thus if we had a used a larger-scale environment, we might have seen differences in the distances of paths walked. Similarly, directional knowledge could be acquired with sufficient exposure to the building, regardless of the interface, consistent with past findings comparing desktop and real-world navigation (Richardson *et al.*, 1999). Visitation errors, in contrast, which involved remembering specific locations and names within the building, were arguably more sensitive measures of learning. This is because subjects not only needed to remember a specific location (a door) but also what name was associated with that door. Indeed, for visitation errors, when participants transferred from immersive VR to the real world we found some evidence for the expected differences as a function of modality: real world > immersive VR > desktop VE. We note, however, that it is possible that if we had used a larger-sized environment, differences would have been evident in normalized path error and directional knowledge as well.

There are two important implications. For building-sized environments, our findings, similar to past studies that compared desktop and real-world navigation (but not immersive VR), suggest that significant spatial learning can indeed take place in VR. Particularly for the accuracy of paths to targets, often taken as a measure of environment-specific knowledge (Newman *et al.*, 2007), our findings demonstrate that subjects can readily acquire such knowledge regardless of the modality. Because we only collected pointing performance on the third and sixth blocks (to avoid disrupting subjects from the VR interfaces to have them point) we cannot say whether transfer might have differed for pointing performance throughout the learning process. Rather, whatever knowledge was needed for the real task could be learned either in the virtual conditions or rapidly acquired in the real world. This also suggests, however, that both immersive VR and desktop VE generally mimic sufficient cues, particularly visual ones, to allow subjects to acquire spatial knowledge regardless of testing modality.

The second implication is that, for arguably more sensitive measures, such as visitation errors, the 'immersiveness' of the modality can impact transfer and knowledge. As discussed previously, we suspect that visitation errors overall required more intimate memory for the spatial layout, and thus the immersive VR interface did confer a slight advantage over the desktop interface for this dependent measure. Similarly, for pointing error, rotational cues in immersive VR would confer an advantage. Notably, however, the real-world condition continued to perform better than the two virtual interfaces. This suggests that real-world navigation continues to capture cues that we could not fully emulate in VR. One reason for this might relate to encoding specificity (Tulving and Thomson, 1973), i.e., that learning will be better for conditions that more fully recapitulate visual and other multisensory cues in the 'encoded' condition compared to those that do not. Although our design did not allow us to specifically investigate this issue, we speculate that visual (i.e., the rendering of depth and other visual cues in VR), somatosensory, and proprioceptive cues, particularly related to movements of the legs and feet, could have potential implications for this difference (Matthis *et al.*, 2018). Another factor might be 'presence' (Kiryu and So, 2007), i.e., the sense that the other VR interfaces do not involve the same rules and contingencies as the real world (e.g., collisions). As VR technology improves and better approximates real-world experiences, future studies will be better able to address this issue as well.

Because of the difficulty of directly comparing virtual and real-world navigation, our setup necessarily involved some limitations. We employed a between-subjects design to try to limit any effect of re-exposure to an environment, although one limitation with between-subjects designs is increased variance due to employing different subjects in each condition. It is possible that a within-subject design could have picked up more subtle differences in the efficiency of walked paths. For immersive VR, we did find increased incidence of cybersickness and dizziness. Possibly, even low levels of discomfort in participants who completed the study could have limited the efficacy of immersive VR, to some extent. This is also a potential drawback of using the immersive VR as a replacement for desktop navigation. Although immersive VR can bring certain benefits to spatial acquisition, some participants might be less likely to finish it. The specific reasons for cybersickness and potential ways to diminish it are therefore important to address in future studies, although we expect such issues to be less of a factor as the technology continues to improve. Finally, our use of a campus building could have potentially obscured differences, particularly in path efficiency, that would have been present during navigation of larger-scale space. Past studies have shown differences in how participants learn different scales of spaces (Meilinger *et al.*, 2016; Montello, 1998; Starrett and Ekstrom, 2018; Starrett *et al.*, 2019). It is quite possible that if participants had to navigate longer distances, such as would be required for a park or city, we would have seen an advantage for immersive over desktop VE.

5. Conclusion

Our results support the view that both immersive VR and desktop navigation are effective spatial learning tools, and, given time, participants readily acquire useable spatial representations in both modalities. Our findings also suggest, however, that immersive technologies confer a slight advantage for when tasking more sensitive measures, like remembering the location of a specific office, although this might be at the cost of an increased dropout.

Acknowledgements

We would like to thank Charlotte Mehaffey for helping to design the virtual Center for Neuroscience. This research was supported by grants from the National Science Foundation (NSF BCS-1630296, PI: ADE) and a Fulbright scholarship (Komise J. W. Fulbrighta, Prague, CZ) to LH.

Supplementary Material

Supplementary material is available online at:
https://doi.org/10.6084/m9.figshare.11611632

References

Brandt, T., Schautzer, F., Hamilton, D. A., Brüning, R., Markowitsch, H. J., Kalla, R., Darlington, C., Smith, P. and Strupp, M. (2005). Vestibular loss causes hippocampal atrophy and impaired spatial memory in humans, *Brain* **128**, 2732–2741.

Chrastil, E. R. and Warren, W. H. (2013). Active and passive spatial learning in human navigation: acquisition of survey knowledge, *J. Exp. Psychol. Learn. Mem. Cogn.* **39**, 1520–1537.

Farrell, M. J., Arnold, P., Pettifer, S., Adams, J., Graham, T. and MacManamon, M. (2003). Transfer of route learning from virtual to real environments, *J. Exp. Psychol. Appl.* **9**, 219–227.

Gallistel, C. R. (1990). *The Organization of Learning. Learning, development, and conceptual change, Vol. viii*. MIT Press, Cambridge, MA, USA.

Grant, S. C. and Magee, L. E. (1998). Contributions of proprioception to navigation in virtual environments, *Hum. Factors* **40**, 489–497.

Kiryu, T. and So, R. H. Y. (2007). Sensation of presence and cybersickness in applications of virtual reality for advanced rehabilitation, *J. Neuroeng. Rehab.* **4**, 34. DOI:10.1186/1743-0003-4-34.

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S. and Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion, *Psychol. Sci.* **9**, 293–298.

Lackner, J. R. and DiZio, P. (2005). Vestibular, proprioceptive, and haptic contributions to spatial orientation, *Annu. Rev. Psychol.* **56**, 115–147.

Loomis, J. M. and Beall, A. C. (1998). Visually controlled locomotion: its dependence on optic flow, three-dimensional space perception, and cognition, *Ecol. Psychol.* **10**, 271–285.

Matthis, J. S., Yates, J. L. and Hayhoe, M. M. (2018). Gaze and the control of foot placement when walking in natural terrain, *Curr. Biol.* **28**, 1224–1233.

Meilinger, T., Strickrodt, M. and Bulthoff, H. H. (2016). Qualitative differences in memory for vista and environmental spaces are caused by opaque borders, not movement or successive presentation, *Cognition* **155**, 77–95.

Montello, D. R. (1998). A new framework for understanding the acquisition of spatial knowledge in large-scale environments, in: *Spatial and Temporal Reasoning in Geographic Information Systems*, M. J. Engenhofer and R. G. Golledge (Eds), pp. 143–154. Oxford University Press, New York, NY, USA.

Montello, D. R., Waller, D., Hegarty, M. and Richardson, A. E. (2004). Spatial memory of real environment, virtual environments, and maps, in: *Human Spatial Memory: Remembering Where*, G. L. Allen (Ed.), pp. 251–285. Lawrence Erlbaum Associates, Mahwah, NJ, USA.

Newman, E. L., Caplan, J. B., Kirschen, M. P., Korolev, I. O., Sekuler, R. and Kahana, M. J. (2007). Learning your way around town: how virtual taxicab drivers learn to use both layout and landmark information, *Cognition* **104**, 231–253.

Richardson, A. E., Montello, D. R. and Hegarty, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments, *Mem. Cogn.* **27**, 741–750.

Ruddle, R. A. and Lessels, S. (2006). For efficient navigational search, humans require full physical movement, but not a rich visual scene, *Psychol. Sci.* **17**, 460–465.

Russell, N. A., Horii, A., Smith, P. F., Darlington, C. L. and Bilkey, D. K. (2003). Long-term effects of permanent vestibular lesions on hippocampal spatial firing, *J. Neurosci.* **23**, 6490–6498.

Starrett, M. J. and Ekstrom, A. D. (2018). Perspective: assessing the flexible acquisition, integration, and deployment of human spatial representations and information, *Front. Hum. Neurosci.* **12**, 281. DOI:10.3389/fnhum.2018.00281.

Starrett, M. J., Stokes, J. D., Huffman, D. J., Ferrer, E. and Ekstrom, A. D. (2019). Learning-dependent evolution of spatial representations in large-scale virtual environments, *J. Exp. Psychol. Learn. Mem. Cogn.* **45**, 497–514.

Taube, J. S., Valerio, S. and Yoder, R. M. (2013). Is navigation in virtual reality with fMRI really navigation?, *J. Cogn. Neurosci.* **25**, 1008–1019.

Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M. and Beall, A. C. (2004). Does the quality of the computer graphics matter when judging distances in visually immersive environments?, *Presence (Camb.)* **13**, 560–571.

Tulving, E. and Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory, *Psychol. Rev.* **80**, 352–373.

Valerio, S. and Taube, J. S. (2012). Path integration: how the head direction signal maintains and corrects spatial orientation, *Nat. Neurosci.* **15**, 1445–1453.

Visell, Y., Giordano, B. L., Millet, G. and Cooperstock, J. R. (2011). Vibration influences haptic perception of surface compliance during walking, *Plos One* **6**, e17697. DOI:10.1371/journal.pone.0017697.

Waller, D., Hunt, E. and Knapp, D. (1998). The tranfer of spatial knowledge in virtual environment training, *Presence (Camb.)* **7**, 129–143.

Waller, D., Loomis, J. M. and Haun, D. B. M. (2004). Body-based senses enhance knowledge of directions in large-scale environments, *Psychon. Bull. Rev.* **11**, 157–163.

Witmer, B. G., Bailey, J. H., Knerr, B. W. and Parsons, K. C. (1996). Virtual spaces and real world places: transfer of route knowledge, *Int. J. Hum. Comput. Stud.* **45**, 413–428.