

Comparing Pedestrian Navigation Methods in Virtual Reality and Real Life

Gian-Luca Savino
University of Bremen
gsavino@uni-bremen.de

Felix A. Kroll
University of Bremen
fe_kr@uni-bremen.de

Rieke Leder
University of Bremen
rleder@uni-bremen.de

Martin Schmeißer
University of Bremen
s_h6krpo@uni-bremen.de

Zihe Xu
University of Bremen
zihe@uni-bremen.de

Jaime Maldonado
University of Bremen
jmalonado@uni-bremen.de

Niklas Emanuel
University of Bremen
emanuel@uni-bremen.de

Marvin C. Lange
University of Bremen
marvin4@uni-bremen.de

Zhanhua Liang
University of Bremen
zhanhua1@uni-bremen.de

Nicolai Schütz
University of Bremen
s_ighm5k@uni-bremen.de

Kerstin Bub
University of Bremen
kerstin.bub@uni-bremen.de

Ernst Kruijff
University of Applied Sciences
Bonn-Rhein-Sieg
ernst.kruijff@h-brs.de

Steven Kowalzik
University of Bremen
ste_kow@uni-bremen.de

Matthis Laudan
University of Bremen
laudan@uni-bremen.de

Dayana Markhabayeva
University of Bremen
dayana1@uni-bremen.de

Carolin Stellmacher
University of Bremen
cstellma@uni-bremen.de

Thorsten Kluss
University of Bremen
tox@uni-bremen.de

Johannes Schöning
University of Bremen
schoening@uni-bremen.de

ABSTRACT

Mobile navigation apps are among the most used mobile applications and are often used as a baseline to evaluate new mobile navigation technologies in field studies. As field studies often introduce external factors that are hard to control for, we investigate how pedestrian navigation methods can be evaluated in virtual reality (VR). We present a study comparing navigation methods in real life (RL) and VR to evaluate if VR environments are a viable alternative to RL environments when it comes to testing these. In a series of studies, participants navigated a real and a virtual environment using a paper map and a navigation app on a smartphone. We measured the differences in navigation performance, task load and spatial knowledge acquisition between RL and VR.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ICMI '19, October 14–18, 2019, Suzhou, China

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6860-5/19/10.

<https://doi.org/10.1145/3340555.3353741>

From these we formulate guidelines for the improvement of pedestrian navigation systems in VR like improved legibility for small screen devices. We furthermore discuss appropriate low-cost and low-space VR-locomotion techniques and discuss more controllable locomotion techniques.

KEYWORDS

Virtual Reality, Navigation, Multi-Modal Interaction

ACM Reference Format:

Gian-Luca Savino, Niklas Emanuel, Steven Kowalzik, Felix A. Kroll, Marvin C. Lange, Matthis Laudan, Rieke Leder, Zhanhua Liang, Dayana Markhabayeva, Martin Schmeißer, Nicolai Schütz, Carolin Stellmacher, Zihe Xu, Kerstin Bub, Thorsten Kluss, Jaime Maldonado, Ernst Kruijff, and Johannes Schöning. 2019. Comparing Pedestrian Navigation Methods in Virtual Reality and Real Life. In *2019 International Conference on Multimodal Interaction (ICMI '19), October 14–18, 2019, Suzhou, China*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3340555.3353741>

1 INTRODUCTION & MOTIVATION

Mobile navigation applications like *Google Maps* are among the most used smartphone apps [14]. Using mobile devices to navigate has become a very frequent task, making mobile map apps the de facto standard for navigation. Still a diverse

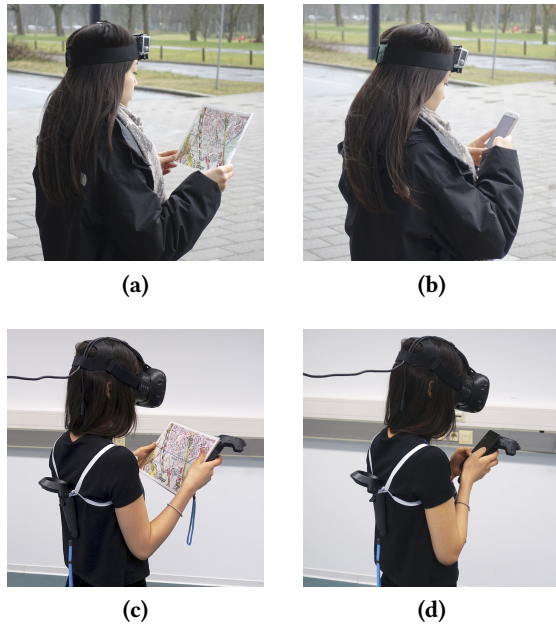


Figure 1: User study conditions: Paper map in RL (a), smartphone in RL (b), paper map attached to a VIVE controller for VR (c), and smartphone attached to a VIVE controller for VR (d).

portfolio of novel navigation methods and multimodal interfaces is continuously developed, such as tactile belts [24], wearable displays [8], drones for displaying spatial information [28] or smartwatch-navigation interfaces [33]. Typically, these methods have been compared against a baseline technique (e.g. paper map or mobile map app) in a field study. However, those field studies in real life environments are hard to compare, as they are typically conducted in different cities with different city structures, different landmarks and different weather and traffic conditions. Therefore, the goal of this paper is to evaluate if VR environments are a viable alternative to RL environments when evaluating navigation interfaces and methods. We highlight problems and challenges of this approach and formulate guidelines on how to evaluate navigation methods in VR and RL. Our work builds upon a large body of research on evaluating navigation methods and devices in the lab [6, 12, 17] and especially in VR [13, 26, 31]. We extend this literature by comparing the effectiveness of this approach and discuss the influence VR has on the two most common navigation methods (paper map and smartphone). If VR environments were to mimic most of the qualities of a RL study, we could potentially eliminate external factors common in RL field studies. A first step towards this is testing the two most common navigation methods under similar conditions in VR and RL to answer

the following research questions: (RQ1) How does the locomotion, the VR environment and the navigation method influence participants navigation performance? And (RQ2) how can we improve upon these factors and bring VR navigation and RL navigation closer together? We created a VR environment that is geographically identical to Findorff a part of Bremen in Germany and compared the navigation performance using a paper map and a smartphone with a mobile map app. Participants performed the same navigation tasks in RL and VR (see figure 1). To test participants in both environments the navigation performance (defined by time and number of errors), task load and the spatial knowledge acquisition were measured. To assure a natural and immersive VR experience we evaluated four different VR locomotion techniques and compared them to walking in RL prior to the main study.

Contributions

Overall, our results show that participants navigating in VR show a different performance from participants in the RL environment. However, as related literature suggests we see promising results for the spatial knowledge acquisition. Therefore we provide insights and formulate guidelines for (but not limited to) the following findings: In VR participants performed better with the smartphone than with the paper map. In RL participants performed better with the paper map than with the smartphone. Individual differences increase in VR. Participants experience higher task load in VR than in RL regardless of the navigation method. Participants acquire the same spatial knowledge in VR and RL in a route recall test. In VR participants quickly adapt to the virtual setting and try to exploit the VR locomotion technique.

2 RELATED WORK

There is a large body of related work studying various aspects of pedestrian navigation. Researchers are interested in navigation strategies and performance [24], task load or spatial knowledge acquisition during navigation [1, 5, 30, 31]. There is also a long tradition of building novel and multimodal navigation prototypes [8, 27] and comparing them to existing techniques such as turn-by-turn navigation on a mobile device or paper map, using field studies. But evaluating navigation tasks in laboratory settings is also becoming more common. Virtual environments (VEs) with different levels of immersion and locomotion techniques have been a widely used tool in those studies. Cliburn [6] for example used three large projection screens oriented at 120 degree angles to conduct an experiment exploring the use of dynamically placed landmarks as navigation aids. Subjects used a gamepad to move and rotate through the VE and interact with the system. For pedestrian street crossing simulations, Deb et al. [10] and Feldstein et al. [13] conducted studies

in VR using low cost HMDs with real walking in a small tracking space. Large screens, projections or HMDs are ways to improve the visual immersion [10, 13, 15, 20, 21] that before was restricted to regular computer screens. Nevertheless there was already extensive research being conducted on navigation that used the benefit of a lab environment even with basic computer monitors.

Spatial Learning in Virtual Environments

Cushman et al. [7] relied on the flexibility and safety of the laboratory in order to investigate whether navigational deficits caused by cognitive ageing and Alzheimer's disease observed in RL could be observed in VR navigation. For this purpose, navigation in RL and VR was compared. In the RL experiment, participants sat in a wheelchair and were carried by an experimenter through an indoor environment following a fixed route. In VR participants were sitting in front of a display and navigated through the simulated environment. Their results show that there is a close correlation between RL and VR navigational deficits and they conclude that virtual navigation is well suited for conducting experiments in this domain. Research also shows that virtual environments can successfully be used to learn spatial information. Wallet et al. [31] have shown that spatial knowledge can be transferred successfully from a VE to the real world when using action based learning. Participants navigated two routes with three different display modes: (1) passive VE (with a route recorded), (2) active VE with joystick and (3) the real environment (the participant actually travelled the route by following instructions). The joystick outperformed the passive condition in both way-finding and sketch mapping tasks. This finding is supported by the results of Richardson et al. [26]. They compared the spatial knowledge acquisition of maps, real- and virtual-indoor environments by letting the participants learn in one of these three conditions and testing their route distance and direction estimates to specific landmarks in the real environment afterwards. Their results point out that all three learning conditions performed similarly on remembering the layout of landmarks on a single floor, though the VE learners achieved the lowest performance. A similar result was found by Bliss et al. by showing that firefighters learning to navigate through a building in VR were almost as successful as the one learning in RL [3]. The above literature shows that lab studies on navigation have been successfully conducted using everything from computer monitors to HMDs and that knowledge and experiences transfer partially from virtual to real environments. Based on this premise we designed our pedestrian simulator in VR and test it against RL conditions.

3 PRE-STUDY: VR LOCOMOTION TECHNIQUES

An existing issue of studies using virtual environments is the break of realism and immersion due to VR locomotion techniques [11]. Therefore, the goal of the preliminary study was to evaluate different locomotion techniques in terms of pace, naturalness and comfort. Unrealistic VR motion disturbs the experience in terms of naturalness and immersion, while users can also suffer from severe motion sickness [2] when exposed to visual-proprioceptive conflicts. A reasonable solution is to match the tracking space to the size of the VE, but is often not feasible because VEs are larger in size, as compared to the space available in laboratories. When the size of the VE exceeds the available space in the laboratory, people have used various complex (hardware) setups to enable locomotion in VR. Examples are the *VirtuSphere* [18], *Suspended Walking* [32], *Omnidirectional Treadmills* [9] or *Redirected Walking* [29]. While those setups offer somehow natural locomotion, they make results hard to compare across different user studies. Therefore we focused on comparing locomotion techniques which do not require additional hardware and have low space and low cost requirements to ensure a wide range of applications and thus comparability. In our pre-study we compare walking in a real environment with *100 ms Warp*, *Avatar Warp*, *Walking-in-Place* and *Freeze Rotation* in a virtual environment, which come at a low price and need a minimal amount of space (3m x 3m). WARP is a sub-type of the arch teleportation location technique. Arch teleportation is the standard locomotion technique in the *Steam VR* home app. AVATAR is also a sub-type of arch teleportation. It requires the participants to choose a target destination, which is then approached by a human avatar. When confirming, the participant is teleported to the avatars position. Walking-in-Place (WIP) measures the head elevation during a stepping motion. The movement is then translated into VR motion, which enables the participant to navigate through the VE while actually walking in place. Freeze Rotation (FREEZE) is a locomotion technique in which participants walk through the VR tracking space. When reaching the border of the tracking space participants can confirm so using a button press which freezes the VR scene. They can then turn around in the tracking space and release the button. This way they have "frozen" the scene and therefore not changed the orientation in VR but can now walk through the VR tracking space again.

Participants & Procedure

We recruited 16 participants (11 male and 5 female) with an average age of 25 (min: 20, max: 30). Every participant navigated a 440m path in RL and VR with the different locomotion techniques. The study used a within-subject design and the conditions were counterbalanced. The participants

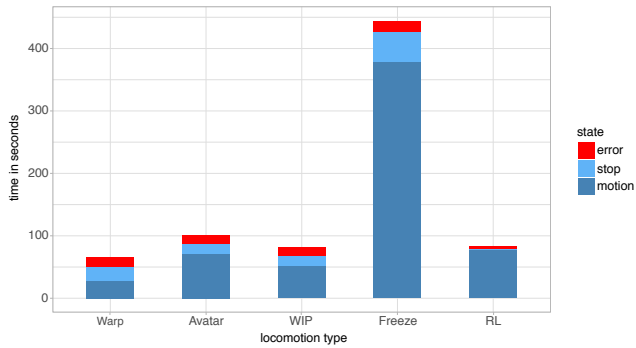


Figure 2: The average time spent in motion, stopping + standing and error correction per segment (110m).

were provided with instructional paper cards that showed the navigation instructions (“left”, “right” or “straight”) at the 24 decision points. The participants were asked to follow the instruction each time they encountered a decision point simulating a turn-by-turn style navigation approach. In RL an experimenter followed them to measure the duration, count the navigation errors and number of stops they made.

Results & Analysis

We conducted a one-way repeated measure analysis of variance (ANOVA) for completion time measurements. It revealed a significant effect of the locomotion technique on completion time performance ($F(4,75) = 186.4, p < .001$). For post-hoc testing we used Sidak-Correction. Pairwise t-testing with repeated measures of the completion time revealed significant differences between FREEZE ($M = 444.65, SD = 95.30$) and all other locomotion techniques respectively. Additionally, WARP ($M = 65.06, SD = 29.73$) and AVATAR ($M = 101.03, SD = 20.49$) showed significant differences ($t(15) = -5.83, p < .001$). When comparing the means WIP ($M = 80.98, SD = 27.98$) was the closest to RL ($M = 83.27, SD = 12.51$), while WARP, AVATAR as well as FREEZE showed larger differences. Each participant made at least 1 navigation error and more than 8 stops with an average of 2.44 errors ($SD = 1.09$) and 16 stops ($SD = 7.42$). The main reasons for stopping and standing were orientation and instruction comprehension. Pairwise Wilcoxon-Signed-Rank tests showed that participants made significantly more stops when moving in each VR locomotion technique compared to RL ($M = 0.45, SD = 0.44$). In terms of the times spent for resolving errors we found no significant effect between the locomotion techniques. In contrast to the time participants stopped, a non-parametric Friedman-Test showed a high significant impact ($\text{Chi-Square} = 39.25, p < .001$) of the locomotion technique on the stop time. In all VR locomotion techniques participants spent significantly more time standing when compared to RL. From all VR locomotion techniques WIP has

the lowest stop rate, standing time and error resolving time. Answers regarding the perceived naturalness and comfort for each locomotion technique were collected using a 6-level forced-choice Likert-scale (0 = not close to RL walking, 5 = very close to RL walking). Post-hoc pairwise Wilcoxon-Signed-Rank tests showed that WIP ($M = 3.19, SD = 0.98$) was significantly higher rated in terms of naturalness when compared to all other locomotion techniques respectively. WARP ($M = 0.94, SD = 1.29$) was rated the least natural, followed by AVATAR ($M = 1.50, SD = 1.21$) and FREEZE ($M = 1.69, SD = 1.62$). A one-way repeated measures ANOVA on the responses about comfort also showed a significant effect between the locomotion techniques ($F(3,60) = 6.01, p = .001$). Post-hoc pairwise t-tests with repeated measures revealed that FREEZE ($M = 1.50, SD = 1.55$) has been considered significantly less comfortable by participants when compared to all other locomotion techniques. The most comfortable locomotion technique was WARP ($M = 3.31, SD = 1.49$), which was slightly better than WIP ($M = 3.00, SD = 1.10$) and AVATAR ($M = 2.94, SD = 1.06$).

We asked participants to rank the perceived accuracy of the pace for each locomotion technique compared to RL walking on a 5 level Likert-scale (much slower, slower, accurate, faster, much faster). Medians for each locomotion technique are as follows: WARP (much faster), Avatar (accurate), WIP (accurate), FREEZE (much slower). When asked to rank the locomotion techniques according to their own assessment half of the participants ranked WARP the best, while 5 participants preferred WIP.

Discussion & Implications

WIP had the closest mean completion time compared to RL walking (see figure 2). In terms of naturalness WIP performed significantly better than the other locomotion techniques, which might be an effect of a more comprehensive full body motion involvement. The only other locomotion technique including a similar body movement was FREEZE, which performed poorly in pace and comfort. Although WARP had the highest level of comfort, it was significantly faster than RL, participants considered it very unnatural and they spent a relatively long time standing and resolving errors (58% of the total time). This highlights the effect of unrealistic and comfortable movement on effective and natural navigation. In both WARP and AVATAR, participants sometimes simply missed a decision point, because they teleported too far, resulting in navigation errors. WIP and FREEZE offered a more continuous motion, but with FREEZE being unacceptably uncomfortable. Therefore we decided to use WIP in the main study as it provides a good balance between pace, naturalness and comfort.

4 MAIN STUDY: VR ENVIRONMENT & NAVIGATION METHODS

The main study investigated participants navigation behaviour with two different methods: A paper map and a mobile map application on a smartphone. The aim was to test these under similar conditions in VR and RL to answer our research questions. (RQ1) How does the locomotion, the VR environment and the navigation method influence participants navigation performance? And (RQ2) how can we improve upon these factors and bring VR navigation and RL navigation closer together?

The study was conducted using a within-subject design meaning each participant tested every navigation method. This results in the following experiment conditions: Navigating in RL with a paper map (RL PM), navigating in RL with a smartphone (RL SP), navigating in VR with a paper map (VR PM) and navigating in VR with a smartphone (VR SP). In each condition the participants performed the navigation task of following a predefined route. Instructions for this were provided by displaying the route on both the paper map and the smartphone respectively in RL and VR. The smartphone additionally provided participants with their current location. To avoid any spatial learning effects between the navigation tasks (conditions) we chose four different routes, which are located in a residential district and randomised the conditions. The area consists of many intersecting streets and allowed us to determine four distinct routes (approx. 420m per route), each of which had five turns.

Measures

This section presents the measures that were used to evaluate the navigation performance, task load, and spatial knowledge in the experiment.

The navigation performance indicates how well participants performed according to the time they took to complete the navigation task and how many errors they made. We call the overall time participants took to complete the route, completion time (CT). It consists of the movement time (MT, time participants are actively moving) and the stop time (ST, time participants stop and stand still). When participants performed a wrong turn at a decision point, an error was counted. The error time (from performing the error until participants were back on the predefined track) was removed from the CT for all conditions.

The task load index is calculated according to the NASA TLX Paper and Pencil Package [16].

The acquired spatial knowledge is measured by two different tests after each condition: A landmark knowledge test and a route knowledge test. Acquired landmark knowledge is determined by a landmark recall task based on the scene recognition test [19]. For each route nine pictures are shown

to the participants of which five are actually showing parts of the route, while the others are unrelated. We ask the participants to decide for each picture if they recall seeing it during the navigation task. For each correct answer they score one point, resulting in a maximum score of nine points per test. The route knowledge test includes a sketch-drawing task that requires the participant to recall the travelled route [31]. We ask the participants to indicate directional changes and the amount of turns by drawing the route on a blank piece of paper.

Implementation

In the RL environment participants received either a laminated DIN A4 sized paper map (figure 1 (a)) or a smartphone (iPhone 7, figure 1 (b)) showing the experiment area and one of the four predefined routes. The iOS app used the *Google Maps* API for a basic map visualisation and to provide the participants with the route information and their current location. The smartphone was also used in the RL PM condition to track navigational data such as the users' location and speed data. For the VR setup we used the *HTC Vive* and the *SteamVR* software. To accurately rebuild the RL environment in VR the layout is based on *OpenStreetMap* data and modelled in *3DS Max*. After completing the basic setup of the environment the geometry was manually refined based on *Google Street View* imagery. Landmarks, street signs and other details were modelled in *Blender* and added to the scene. We use a slightly modified version of the WIP technique compared to the one used in the preliminary study. This version uses one of the VIVE controllers worn as a backpack (as depicted in figure 1 (c & d)). The controller's up and down movement was tracked to translate the participant's WIP motion into a forward motion in VR. The grip button of the other controller was used to toggle the usage of the WIP locomotion. To simulate the natural haptic feedback of the paper map and the smartphone they were attached to the controller with velcro (see figure 1 (c & d)), depending on the condition. To match the participants preferred hand the controller could be attached on either the map's right or left side. The smartphone ran a network client application created in *Unity* which allowed us to use the full multi-touch functionality of the device. The application used *UNET*, *Unity's* high-level networking API to communicate with the main study setup. All input was evaluated on the smartphone, while the detected gestures (e.g. touch, pan, pinch and rotate) were sent to the server. The corresponding actions were then performed on the virtual smartphone, making for a fully immersive smartphone interaction.

Participants & Procedure

We recruited 16 participants (8 male and 8 female) with an average age of 24 (min: 18, max: 30) through advertisements

on local online bulletin boards. People with strong visual or hand-motor system impairments were excluded. In a demographic survey 11 participants stated to have used HMDs for a few times, 4 participants never used HMDs before and one participant has been using it on a regular basis. All participants use navigation applications, varying from every day to rarely. Depending on the starting condition for the participants, we met them either at the RL environment or at the laboratory where the VR setup was located. In RL they were guided to the respective starting point of each route. At the starting point participants received instructions for the navigation task, a remark about the tests after each task and were equipped with a *GoPro* (see figure 1 (a & b)). Participants performed the task on their own and were picked up at the route's destination. After reaching the destination, the participants performed the *NASA TLX* and landmark recall test on an iPad, the route recall test was done with pen and paper. The VR part of the experiment started with a three minute introduction to the *HTC VIVE* and a training for the WIP locomotion technique. Afterwards the participants were introduced to the first navigation method and had some additional time to get used to it. Afterwards they were placed into the experiment environment and performed the task. Like in the RL condition the participants had to take a route and landmark recall test as well as the *NASA TLX* after each run. After the experiment they were asked to fill in a demographic questionnaire also including questions about their VR experience (e.g. if they experienced simulator sickness) during the experiment. This was followed by a semi-structured interview. The interview was done in the participants' preferred language, either English or German and included questions about their perceived difference between RL and VR as well as between the navigation methods. Between the VR and RL conditions participants were escorted by car (10 minutes) to their next destination.

5 RESULTS & ANALYSIS

Half of the participants reported having simulator sickness in VR and rated the influence on their performance with 2.5 on average (0 = no influence, 4 = strong influence). When asked about how familiar participants were with the RL environment the average answer was 2.25 (0 = not familiar at all, 5 = very familiar).

Navigation Performance

In RL the average completion time was 266.38 seconds ($SD = 32.30$) when using a paper map and 287.12 seconds ($SD = 44.87$) when using a smartphone. In VR the participants were faster in both the PM ($M = 230.23$, $SD = 100.67$) and SP condition ($M = 195.03$, $SD = 102.40$). The standard deviation was 2-3 times larger in VR than in RL, which indicates higher individual differences in VR. A Wilcoxon-Signed-Rank test

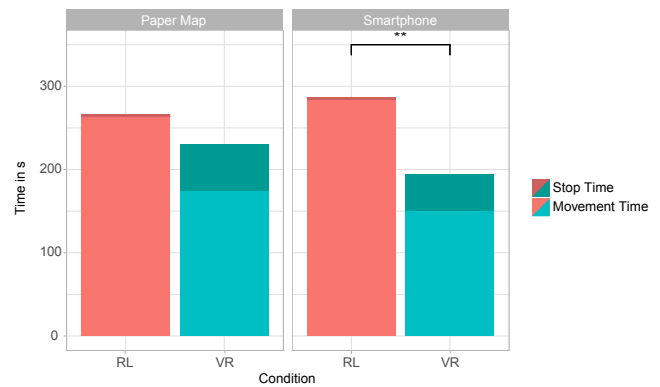


Figure 3: Average completion time for each method in each condition.

revealed significant differences when comparing the SP conditions ($p < .05$, $Z = -2.543$), whereas no significance was found in the PM condition ($p = 0.193$, $Z = -1.302$). Figure 3 shows the mean CTs for all conditions.

Movement Time. When using the paper map participants were on average slower in RL ($M = 263.16$, $SD = 35.86$) than in VR ($M = 173.97$, $SD = 91.15$). Similar results were found in the smartphone conditions, where in RL ($M = 283.76$, $SD = 43.88$) the movement time was higher than in VR ($M = 150.78$, $SD = 79.04$). Pairwise t-tests with repeated measures of the movement time revealed significant differences for both conditions (PM: $t(15) = 4.26$, $p < .001$, SP: $t(15) = 5.81$, $p < .0001$). However, in RL the movement time when using a smartphone was increased compared to the paper map condition, while in VR we found the opposite.

Stop Time. When using the paper map in RL 4 participants stopped at least once, while the others walked continuously. The smartphone induced slightly more stops, but still five participants never stopped. In VR all participants stopped at least once, regardless of the method. Correlating to these numbers was the difference in stop time between RL and VR. A Wilcoxon-Signed-Rank test found significant differences ($p < .001$, $Z = -3.491$) when comparing the stop time in RL PM ($M = 3.22$, $SD = 6.99$) to the VR counterpart ($M = 56.26$, $SD = 37.65$). Performing the same test on the smartphone conditions also revealed significant differences ($p < .0001$, $Z = -4.170$) between RL ($M = 3.36$, $SD = 3.54$) and VR ($M = 44.25$, $SD = 30.16$). Considering the standard deviation our results suggest notably large individual differences in all conditions.

Errors. In total 30 navigation errors were recorded. 7 (PM: 2, SP: 5) errors were performed in RL and 23 (PM: 10, SP: 13) in VR. When comparing the navigation methods in both RL and VR the smartphone induced more errors than the

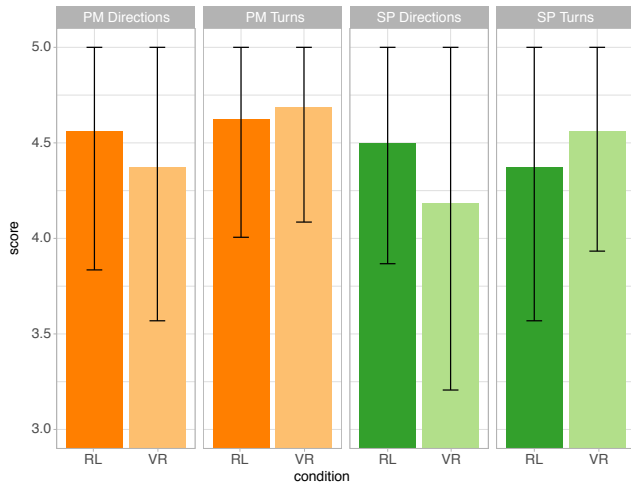


Figure 4: Average route recall score for each condition in terms of number of correct turns and number of correct directional changes.

paper map. It is important to note that three participants were responsible for 24 (80%) of all errors which include all of the RL errors and 17 of the VR errors. All remaining 6 errors were performed in VR by different participants.

Task Load

When using the paper map 14 out of 16 participants considered the task load in RL ($M = 30.33$, $SD = 20.77$) to be lower than in VR ($M = 47.25$, $SD = 18.73$). In the smartphone condition all of the participants reported a lower task load in RL ($M = 27.54$, $SD = 16.45$) than in VR ($M = 47.88$, $SD = 18.50$). On average the task load was $\sim 36\%$ less in RL than in VR for the paper map and $\sim 42\%$ for the smartphone condition. A Wilcoxon-Signed-Rank test found significance for both the paper map ($p < .001$, $Z = -3.366$) and smartphone condition ($p < .0001$, $Z = -4.170$). When comparing every single subscale of the task load index between RL and VR we always found a lower mean in RL than in VR, regardless of the method. The results suggest that the task load in RL is generally lower than in VR and that the navigation method itself does not seem to have a large impact.

Spatial Knowledge

The measurements of the acquired spatial knowledge consisted of a landmark recall and a route recall test. In the landmark recall test participants could achieve a maximum score of 9, while the route recall test was divided into two measurements with a maximum score of 5 points each.

Landmark Recall Test. On average the landmark recall was better in RL than in VR for both paper map and smartphone. In the paper map condition the average score decreased

from 5.94 in RL to 5.06 in VR, while the standard deviation increased slightly from 1.48 to 1.77 respectively. We found similar results for the smartphone condition, where the mean score decreased from 5.94 in RL to 4.75 in VR, while the standard deviation remained similar (1.18 in both conditions). A pairwise t-test with repeated measures found no significant differences ($t(15) = -1.28$, $p = 0.2192$) between RL and VR in the paper map condition, but significant differences were revealed when comparing the smartphone conditions ($t(15) = 3.05$, $p < .01$). When comparing the landmarks themselves we found that in RL all landmarks were recognised at least once, while there were two landmarks in VR that were never recalled by the participants. In VR 14 out of 20 landmarks had a recognition rate below 50%, while in RL 10 out of 20 had a recognition rate below 50%.

Route Recall Test. No participant scored less than 3 out of 5 possible points for both the number of correct turns ($\#turns$) and the number of correct directional changes ($\#directions$) respectively. The average score in terms of the correct amount of turns slightly increased when comparing RL (PM: 4.63, SP: 4.38) to VR (PM: 4.69, SP: 4.56). In contrast to that the average score for the correct directional changes is fairly higher in RL (PM: 4.56, SP: 4.50) than in VR (PM: 4.38, SP: 4.19). A pairwise Wilcoxon-Signed-Rank test revealed no significant difference for both aspects between RL and VR for paper map ($\#turns$: $p = 0.766$, $Z = -0.298$, $\#directions$: $p = 0.429$, $Z = -0.791$) and smartphone ($\#turns$: $p = 0.407$, $Z = -0.829$, $\#directions$: $p = 0.243$, $Z = -1.167$). Nevertheless it is worth to mention that the smartphone caused a slight decrease in the means with a slightly bigger standard deviation for each aspect in RL and VR compared to the paper map which can be seen in figure 4.

6 DISCUSSION & CONCLUSION

In our study we compare navigation performance, task load and spatial knowledge acquisition using a paper map and smartphone in RL and VR respectively. Overall the results between VR and RL were significantly different in navigation performance, task load and landmark recognition. Route recognition, however, was not significantly different in both RL and VR. Still challenges and problems are to overcome until VR environments can fully replace RL for navigation in an experiment setting. The following discussion will therefore highlight how the locomotion, the VR environment and the navigation method influence participants navigation performance (RQ1). We also present guidelines (see figure 5) on how we can improve upon these factors and bring VR navigation and RL navigation closer together (RQ2).

Issue	Guidelines	Based on
Individual differences in locomotion speed	Individually measure and adapt walking speeds per participant	Recorded data
Increasing locomotion speed over time	Use measured walking speeds as threshold + long training phases (>15min)	Recorded data
Environment for landmark knowledge	Choose/Create environments that include "landmark worthy" architecture/objects	Recorded data
Tracking of real props for virtual devices	Use external tracking like Leap Motion to track hands and devices	Development Phase
Novelty factor of VR	Recruit experienced users	General observation
Narrow FOV of HMDs	New HMDs will come with wider FOV, for now narrowing the FOV in RL is an option	General observation
Legibility of small screen devices or maps	Use latest HMDs for highest display resolution	General observation

Figure 5: Guidelines for testing navigation methods in VR based on different findings of our study.

Locomotion

WIP proved to be the most natural locomotion technique from the ones that were tested in our pre-study, but that doesn't mean that it doesn't come with drawbacks that need to be discussed. McMahan's Framework for Interaction Fidelity Analysis (FIFA) rates WIP slightly above mid-fidelity and found that mid-fidelity interactions mostly perform worse than high- as well as low-fidelity interaction [22, 23]. Although their implementation differed slightly from ours we would rate our implementation similarly and see the following challenges for pedestrian navigation: The movement times in figure 3 suggest that participants were overall faster in the VR conditions compared to the RL conditions. This could be an artefact of the factor translating the up and down movement into forward motion while using our WIP implementation. To better control this in future studies we suggest taking individual walking speed measure in the beginning of both the VR and RL condition. That way experimenters know how much the speed gets in- or decreased in VR on an individual basis. Another thing we noted was, that participants sped up during the second VR condition as they got used to WIP and tried to exploit the locomotion technique (e.g. jogging, walking through objects). This could be dealt with using the before measured walking speed as a reference and prohibit deviating too much from it.

VR Environment

In comparison between the RL and VR conditions the route recall test did not show significant differences. Due to the short length of the routes it was possible, on both the paper map and the smartphone, to see the whole route at once. So participants could have potentially recalled the image of the route they saw on their device. Due to the high number of landmarks that we used in our test the landmarks chosen in RL and remodelled in VR, some of them might not have been recognised as such. As most of the streets were regular neighbourhoods we had to deal with a lot of repeating

patterns in architecture which made it difficult to generate enough building that qualify as proper landmarks.

As the questionnaire showed, only one participant used VR technology regularly. The rest of the participants have used it either a few times or never before. Being exposed to new technologies often induces a high task load. As our implementation of the WIP technique and the use of our VR props was unique in this setup we expected this to have an influence on the NASA TLX results.

As mentioned in the results we saw a high number of errors were made by only 3 participants. Two of these three experienced strong simulator sickness. However we didn't find any significant correlation in VR between the errors and simulator sickness or errors and familiarity with the environment.

Navigation Methods

The 110 degrees field of view (FOV) and the low resolution compared to natural vision are the top confounding factors when using a HMD in VR experiments. We found that less FOV results in increased head movement as participants look around or at their navigation device. This potentially decreases the spatial knowledge acquisition and increases the physical demand. The low resolution introduces legibility issues when reading small text on a map or smartphone. Participants had to bring the smartphone screen unnaturally close to their face to read street names when comparing them to the street signs in the VR environment. Especially for smartphone navigation in virtual city environments HMDs with higher display resolution will improve the benefit of street names and other indicators on maps. Most VR implementations use bigger text placed directly in the environment to ensure legibility [4] but due to the realistic form factor of our navigation devices we were bound to a certain maximum text size. The results show that in RL participants were a slightly faster with the paper map than with the smartphone when in VR it was the other way round. Additionally in RL participants performed more than double the numbers of errors with the smartphone compared to the paper map whereas in

VR the score was a lot more similar. The RL difference could be explained by the simplicity and length of the route. With a longer, more complex route the smartphone might have had the upper hand due to the possibility to zoom into the area of interest but with a short route the map gave a bigger more readable overview in RL. In VR however participants were faster with the smartphone. As figure 1 shows participants used real props with the *HTC Vive* controller attached which they could manipulate in VR. This made for a very realistic experience. However it worked better for the smartphone than for the paper map, which could explain the difference in CT and the larger number of errors for the paper map. The controller attached to the smartphone served as a handle to hold the phone and was evenly balanced. The controller attached to the map created an imbalance as now one side of the map was heavier than the other which hindered turning the map to face track-up, which is a common behaviour while navigating [25]. Tracking the map and the hands with external hardware could be a solution to resolve this issue. On the smartphone rotation could be easily be performed in VR by using the well known two finger rotate-gesture. According to McMahan's FIFA [22] our VR paper map and smartphone interaction rates as high-fidelity which is reflected in our general observation of participants using them, especially for the smartphone. The overall touch input on the smartphone seemed to work very well for participants. In prior tests we found that touch accuracy is very high even without highlighting the touch points in VR, as long as the VR model and the RL prop are perfectly aligned. Having the controller attached to the phone in VR also helped a lot. Interestingly the task load was not affected by the navigation method.

VR offers a promising simulation environment to test navigation techniques, but comes with lots of challenges that prevent a true to RL navigation experience. Our research shows that the interaction with two of the most common navigation aids can be successfully implemented using a VE and HMDs in a pedestrian navigation simulation in VR. Still typical problems like field of view and legibility are yet to be solved but will lose importance as the technology develops. Especially the latter has a large influence on the usability of text based navigation devices in VR. We believe that our results and guidelines will help future researchers to compare multi modal navigation interfaces in VR as well as motivate future work to close the gap between RL field studies and VR lab studies bit by bit.

REFERENCES

- [1] İlhan Aslan, Maximilian Schwalm, Jörg Baus, Antonio Krüger, and Tim Schwartz. 2006. Acquisition of Spatial Knowledge in Location Aware Mobile Pedestrian Navigation Systems. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '06)*. ACM, New York, NY, USA, 105–108. <https://doi.org/10.1145/1152215.1152237>
- [2] Willem Bles, Jelte E Bos, Bernd de Graaf, Eric Groen, and Alexander H Wertheim. 1998. Motion sickness: only one provocative conflict? *Brain Research Bulletin* 47, 5 (1998), 481 – 487. [https://doi.org/10.1016/S0361-9230\(98\)00115-4](https://doi.org/10.1016/S0361-9230(98)00115-4)
- [3] James P. Bliss, Philip D. Tidwell, and Michael A. Guest. 1997. The Effectiveness of Virtual Reality for Administering Spatial Navigation Training to Firefighters. *Presence: Teleoper. Virtual Environ.* 6, 1 (Feb. 1997), 73–86. <https://doi.org/10.1162/pres.1997.6.1.73>
- [4] Doug A. Bowman, Chris North, Jian Chen, Nicholas F. Polys, Pardha S. Pyla, and Umur Yilmaz. 2003. Information-rich Virtual Environments: Theory, Tools, and Research Agenda. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '03)*. ACM, New York, NY, USA, 81–90. <https://doi.org/10.1145/1008653.1008669>
- [5] Stefano Burigat and Luca Chittaro. 2011. Pedestrian Navigation with Degraded GPS Signal: Investigating the Effects of Visualizing Position Uncertainty. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 221–230. <https://doi.org/10.1145/2037373.2037407>
- [6] Daniel Cliburn, Tess Winlock, Stacy Rilea, and Matt Van Donsel. 2007. Dynamic Landmark Placement As a Navigation Aid in Virtual Worlds. In *Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology (VRST '07)*. ACM, New York, NY, USA, 211–214. <https://doi.org/10.1145/1315184.1315225>
- [7] L. A. Cushman, K. Stein, and C. J. Duffy. 2008. Detecting navigational deficits in cognitive aging and Alzheimer disease using virtual reality. *Neurology* 71, 12 (Sep 2008), 888–895.
- [8] Alexandru Dancu, Mickaël Fourgeaud, Mohammad Obaid, Morten Fjeld, and Niklas Elmqvist. 2015. Map Navigation Using a Wearable Mid-air Display. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 71–76. <https://doi.org/10.1145/2785830.2785876>
- [9] Rudolph P. Darken, William R. Cockayne, and David Carmein. 1997. The Omni-directional Treadmill: A Locomotion Device for Virtual Worlds. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology (UIST '97)*. ACM, New York, NY, USA, 213–221. <https://doi.org/10.1145/263407.263550>
- [10] Shuchisnigdha Deb, Daniel W. Carruth, Richard Sween, Lesley Strawderman, and Teena M. Garrison. 2017. Efficacy of virtual reality in pedestrian safety research. *Applied Ergonomics* 65 (2017), 449 – 460. <https://doi.org/10.1016/j.apergo.2017.03.007>
- [11] Ioannis Delikostidis, Holger Fritze, Thore Fechner, and Christian Kray. 2015. *Bridging the Gap Between Field- and Lab-Based User Studies for Location-Based Services*. Springer International Publishing, Cham, 257–271. https://doi.org/10.1007/978-3-319-11879-6_18
- [12] Martin J. Farrel, Paul Arnold, Steve Pettifer, Jessica Adams, Tom Graham, and Michael MacManamon. 2003. Transfer of Route Learning From Virtual to Real Environments. *Journal of Experimental Psychology: Applied* 9, 4 (2003), 219–227. <https://doi.org/10.1037/1076-898X.9.4.219> PMID: 14664673.
- [13] Ilja Feldstein, André Dietrich, Sasha Milinkovic, and Klaus Bengler. 2016. A Pedestrian Simulator for Urban Crossing Scenarios. *IFAC-PapersOnLine* 49, 19 (2016), 239 – 244. <https://doi.org/10.1016/j.ifacol.2016.10.531> 13th IFAC Symposium on Analysis, Design, and Evaluation of Human-Machine Systems HMS 2016.
- [14] Dan Frommer. 2017. These are the 10 most popular mobile apps in America. Blog. Retrieved August 28, 2017 from <http://www.recode.net/2017/8/24/16197218/top-10-mobile-apps-2017-comscore-chart-facebook-google>.

- [15] Simpson Gordon, Johnston Lucy, and Richardson Michael. 2003. An investigation of road crossing in a virtual environment. *Accident Analysis & Prevention* 35, 5 (2003), 787 – 796. [https://doi.org/10.1016/S0001-4575\(02\)00081-7](https://doi.org/10.1016/S0001-4575(02)00081-7)
- [16] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139 – 183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [17] Jesper Kjeldskov, Mikael B. Skov, Benedikte S. Als, and Rune T. Høegh. 2004. Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. In *Mobile Human-Computer Interaction - MobileHCI 2004*, Stephen Brewster and Mark Dunlop (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg.
- [18] Thorsten Kluss, William E. Marsh, Christoph Zetsche, and Kerstin Schill. 2015. Representation of impossible worlds in the cognitive map. *Cognitive Processing* 16, 1 (01 Sep 2015), 271–276. <https://doi.org/10.1007/s10339-015-0705-x>
- [19] Brigitte Lapeyre, Sylvain Hourlier, Xavier Servantie, Bernard N’Kaoua, and Hélène Sauzèon. 2011. Using the Landmark-Route-Survey Framework to Evaluate Spatial Knowledge Obtained From Synthetic Vision Systems. *Human Factors* 53, 6 (2011), 647–661. <https://doi.org/10.1177/0018720811421171>
- [20] Christian Lehsing, Ilja Feldstein, André Dietrich, and Klaus Bengler. 2016. Pedestrian Simulator for Traffic Research - State of the Art and Future of a Motion Lab. (06 2016).
- [21] Luigi Maffei, Massimiliano Masullo, Francesco Sorrentino, and Maria Gabriele. 2014. Preliminary studies on the relation between the audio-visual cues’ perception and the approaching speed of electric vehicles. 20 (04 2014), 1–9.
- [22] Ryan P. McMahan. 2011. *Exploring the Effects of Higher-Fidelity Display and Interaction for Virtual Reality Games*. PhD dissertation. Virginia Tech.
- [23] Ryan P. McMahan, Chengyuan Lai, and Swaroop K. Pal. 2016. Interaction Fidelity: The Uncanny Valley of Virtual Reality Interactions. In *Virtual, Augmented and Mixed Reality - 8th International Conference, VAMR 2016, Held as Part of HCI International 2016, Toronto, Canada, July 17-22, 2016. Proceedings*. 59–70. https://doi.org/10.1007/978-3-319-39907-2_6
- [24] Martin Pielot and Susanne Boll. 2010. Tactile Wayfinder: Comparison of Tactile Waypoint Navigation with Commercial Pedestrian Navigation Systems. In *Pervasive Computing*, Patrik Floréen, Antonio Krüger, and Mirjana Spasojevic (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 76–93.
- [25] Daniel R Montello and Corina Sas. 2006. Human Factors of Wayfinding in Navigation. *International Encyclopedia of Ergonomics and Human Factors* (03 2006). <https://doi.org/10.1201/9780849375477.ch394>
- [26] Anthony E. Richardson, Daniel R. Montello, and Mary Hegarty. 1999. Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition* 27, 4 (01 Jul 1999), 741–750. <https://doi.org/10.3758/BF03211566>
- [27] Maximilian Schirmer, Johannes Hartmann, Sven Bertel, and Florian Ehtler. 2015. Shoe Me the Way: A Shoe-Based Tactile Interface for Eyes-Free Urban Navigation. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 327–336. <https://doi.org/10.1145/2785830.2785832>
- [28] Stefan Schneegeass, Florian Alt, Jürgen Scheible, Albrecht Schmidt, and Haifeng Su. 2014. Midair Displays: Exploring the Concept of Free-floating Public Displays. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 2035–2040. <https://doi.org/10.1145/2559206.2581190>
- [29] Evan A. Suma, Mahdi Azmandian, Timofey Grechkin, Thai Phan, and Mark Bolas. 2015. Making Small Spaces Feel Large: Infinite Walking in Virtual Reality. In *ACM SIGGRAPH 2015 Emerging Technologies (SIGGRAPH '15)*. ACM, New York, NY, USA, Article 16, 1 pages. <https://doi.org/10.1145/2782782.2792496>
- [30] Delphine Szymczak, Kirsten Rassmus-Gröhn, Charlotte Magnusson, and Per-Olof Hedvall. 2012. A Real-world Study of an Audio-tactile Tourist Guide. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '12)*. ACM, New York, NY, USA, 335–344. <https://doi.org/10.1145/2371574.2371627>
- [31] Grégory Wallet, Hélène Sauzèon, Jérôme Rodrigues, and Bernard N’Kaoua. 2008. Use of Virtual Reality for Spatial Knowledge Transfer: Effects of Passive/Active Exploration Mode in Simple and Complex Routes for Three Different Recall Tasks. In *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology (VRST '08)*. ACM, New York, NY, USA, 175–178. <https://doi.org/10.1145/1450579.1450616>
- [32] Benjamin Walther-Franks, Dirk Wenig, Jan Smeddinck, and Rainer Malaka. 2013. Suspended Walking: A Physical Locomotion Interface for Virtual Reality. In *Entertainment Computing – ICEC 2013*, Junia C. Anacleto, Esteban W. G. Clua, Flavio S. Correa da Silva, Sidney Fels, and Hyun S. Yang (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 185–188.
- [33] Dirk Wenig, Johannes Schöning, Brent Hecht, and Rainer Malaka. 2015. StripeMaps: Improving Map-based Pedestrian Navigation for Smartwatches. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 52–62. <https://doi.org/10.1145/2785830.2785862>