Wired to exit: Exploring the effects of wayfinding affordances in underground facilities using Virtual Reality

Abstract

A virtual reality (VR) experiment with twenty-four participants was conducted using a wayfinding installation with the Oculus Rift S HMD. Participants were immersed into a simulation of a burning underground parking lot and tasked to navigate to the exit. The purpose was to investigate the high-level effect of wayfinding assistive lights on behavioral, physiological, and psychological outcomes. Participants were split into two groups: the baseline condition group was exposed to a scene without assistive lights, and the experimental condition group was exposed to the same scene with assistive lights. Results show that participants in the baseline condition traveled more, made more pauses and turns and took more time to find the exit, but these differences were not found to be statistically significant. On the contrary, differences in heart rate (HR) outcomes between the two groups were found to be statistically significant, with subjects in the baseline condition displaying an increasing HR trend during simulation. This finding is aligned with prior results on the efficacy of landmarks and wayfinding affordances in reducing cognitive demands by improving brain wiring efficiency. We discuss these findings in the context of a rich wayfinding affordances literature.

Keywords

Virtual Reality, wayfinding, underground parking, heart rate, Polar A370

Introduction

Emergencies are critical incidents with potential for highly negative outcomes (Feng, González, Amor, Lovreglio & Cabrera-Guerrero, 2018). In addition, human behavior in these incidents can be erratic and unpredictable (Zhao, Ming Lo, Zhang & Liu. 2009). Immersive virtual reality provides a possibility to simulate emergencies including Chemical, Biological, Radiological, Nuclear, and high-yield Explosive (CBRNE) events without causing physical harm. Controlled simulated settings allow people to train for rational behavior in emergencies to improve their chance of staying operational in high risk situations. While simulations for training and therapy are the most widely researched application areas of VR, they still hold many unanswered questions. The emotional impact of simulations and the resulting influence on performance is one of the less unexplored topics, possibly due to the complexity of the required research setups and necessity of mixed methods approaches combining different complementing measurements for comprehensive results.

A virtual reality (VR) experiment with twenty-four participants was conducted using a wayfinding installation with the Oculus Rift S HMD. Participants were immersed into a simulation of a burning underground parking lot and tasked to navigate to the exit. The purpose was to investigate the high-level effect of wayfinding assistive lights on behavioral, physiological, and psychological outcomes. Participants were split into two groups. The baseline condition group was exposed to a scene without assistive lights, while the experimental condition group was exposed to the same scene with assistive lights. Results show that participants in the baseline condition traveled more, made more pauses and turns and took more time to find the exit, but these differences were not found to be statistically significant. However, differences in heart rate (HR) outcomes between the two groups were found to be statistically significant, with subjects in the baseline condition displaying an increasing HR trend during simulation. This finding is aligned with prior results on the efficacy of landmarks and wayfinding affordances in reducing cognitive demands by improving brain wiring efficiency. We discuss these findings in the context of a rich wayfinding affordances literature.

Related work

CBRNE events can destabilize governments, enable conditions that aggravate violence, or promote terrorism. Furthermore, these events can quickly wear down the civil protection infrastructure and capability of the responders - leading to loss of trust in public institutions (Su and Anthony, 2019; van Buuren and Wijnmalen, 2015). Ports are a critical part of the transportation infrastructure, which also makes them critical when planning and designing CBRNE training scenarios (Christopher, 2014). Offshore assets, like ships and oil rigs, are also vulnerable to CBRNE attacks. And in this case, it is particularly important that their crews and first responders at sea (usually navy crews) are trained to deal with this threat, even though that is not their primary job (May, 2004). The research presented in this paper builds off the back of the XXXX project (XXX, XXX), that aims to support front-line emergency medical responders, law enforcement/security services and ship crews, by providing them with Virtual Reality Simulations (VRS). VRs are known to increase the immersion of training scenarios with realistic physiological and emotional responses. The market for commercial VR technology is expected to generate 98.4 million in sales by 2023 and to reach a worldwide population penetration of 2%. Therefore, many public and private organizations across the globe are adapting VR systems for training responders on dangerous situations that cannot be illustrated accurately using traditional media (Rogers, 2019). VR training, featuring evacuation and wayfinding scenarios, can commence in realistic environments, such as a seaport or other sites under construction without disturbing the logistic flow, decision making, and planning (Tang et al., 2009). For example, the European Organization for Nuclear Research recently conducted a feasibility study for the construction of the Future Circular Collider (FCC), that included VR solutions for evacuation wayfinding experiments in large underground tunnels (Arias et al., 2019).

Wayfinding and emergency simulations

Since the 1960's when the term first debuted in the scientific discourse (Lynch, 1960), wayfinding has been used extensively in evacuation and navigation experiments that seek to improve understanding of behavioral outcomes and determine the optimal system-design for enhancing public security and individual safety. There is a vast literature based on navigation experiments on diverse risk scenarios and locales including trail tunnels, road tunnels, indoor locations (Tucker et al. 2018; Zhao et al., 2009), flood events and emergency exit portals to name a few. Moreover, experiments with wayfinding installations often control design parameters like signage systems that include photoluminescent and auditory cues as navigation affordances. Recently, advances in 3D computer graphics and robotics has spurred interest in navigation experiments using Immersive Virtual reality. The VR systems covered by this emerging literature contain contraptions ranging from mobile, desktop, and large displays to highly immersive head-mounted displays (HMD), Fish Tank and Cave Automatic Virtual Environment (CAVE) and panorama manifestation (PM) systems (Ronchi et al. 2019; Ronchi et al. 2016; Meng & Zhang, 2014).

Metrics for VR simulations

Congruently, wayfinding experiments in virtual reality have paved the way for new and interesting questions on how to best capture, evaluate and measure behavioral and psychometric outcomes. Evaluation metrics are largely based on three Key Performance Indicators (KPIs): (1) user's task/behavioral performance, measured with back-end analytics software running in the background during the simulation, (2) physical behavior, tracked with various head and body tracking sensors (Radianti et al, 2020; Feng et al., 2019; Williams-Bell et al. 2015), and (3) self-reported psychometric and reasoning measures, archived with several qualitative protocols including go-along (Kostakos et al. 2019; Feng et al., 2018) and think-aloud interviews and structured questionnaires (Ruddle and Lessels 2006). In regards to the evaluation of behavioral outcomes with in-game analytics, prior studies exploring wayfinding systems in VR have widely acknowledged the significance of at least five key levels of analysis: i) Distance travelled by the user in order to complete the objective/mission of the game; ii) Time required to complete the objective/mission; iii) Pauses during the simulation that are calculated as user inactivity using time (seconds) or graphics (frames) heuristics; iv) Speed of the user, measured as the total distance travelled divided by time spent; and v) total area covered by the user during the simulation.

Besides behavioral indicators, a few studies have used biometric measurements, although under limited laboratory conditions, for characterizing physiological outcomes. For instance, recent systematic reviews on evacuation experiments using VR listed only a handful of studies that analyzed the impact of the simulation on cardiac physiological outcomes (Radianti et al, 2020; Feng et al., 2019; Williams-Bell et al. 2015). Meng and Zhang (2014) developed fire evacuation scenarios with 40 participants in a virtual hotel. The treatment group played a simulation that included virtual fire effects and explosions while the control group played the same simulation but without the pyrotechnic effects. Participants' evacuation time, movement pathway, and heart rate was recorded. The treatment group (simulation with pyrotechnics) had a statistically significant higher mean Heart Rate Ratio (HRR). Furthermore, the study noted that for both groups the HR increased once the simulation had just started - a pattern that seems to be consistently observed across studies in the field (Tucker et al. 2018). While the Meng and Zhang experiment (2014) presents interesting results, the differences in the HR data can be attributed to the intense auditory and visual artifacts and thus, do not enable direct links with the design of the signage system per se. Correspondingly, (Chittaro and Buttussi, 2015) tested a VR game that simulates a realistic emergency landing and evacuation experience in a commercial aircraft. Participants were fitted with a raw blood volume pulse sensor (BVP) to index physiological arousal. The BVP amplitude was also recorded. They found substantial differences between the group that played the VR simulation and the baseline group that only read evacuation instructions from a non-VR safety card. Like Meng and Zhang's (2014) study, variations in the HR readings cannot be attributed to design decisions since the control group was not exposed to the same VR experience as the other group.

Thus, the question remains whether wayfinding affordances and signs within a VR simulation – all other things being equal— can have a significant effect on physiological outcomes as captured through biometric sensors. This is a key question for improving our understanding on how to best control the levels of stress and anxiety in evacuation simulations, as well as in evaluating the effects of wayfinding affordances and signs under demanding conditions.

The paper presents results of an experiment that explores behavioral and physiological outcomes of users performing wayfinding tasks with and without photoluminescent landmarks. Our experimental set-up consists of a commercially available head-mounted display which is used to display an evacuation scenario in a custom-designed 3D underground parking space. To overcome the limitations of past studies, we introduce the use of a photoplethysmogram (PPG) sensor for monitoring changes in heart rate (HR) levels. The widespread availability of wearable and cost-effective biosensors (EEG and PPG) is leading to the increased use of these devices in research settings to monitor cognitive workloads and to enable brain-computer interfaces in VR

(Feehan et al. 2018; Henriksen et al. 2018). This scenario was chosen mainly because it allows the use of a safe and controllable environment and is realistic for both offshore and onshore facilities. Later tests will be conducted with professionally trained experts at the Portuguese Navy Damage Control School facilities, that include a pier, a small village, and other buildings.

Based on our results, we can suggest an optimal setup for training emergency respondents for working in confined indoor spaces with low visibility, such as silos, storage tanks in the harbor areas or in particular, aboard ships, cruise vessels and ferries. Our main goal is to use what we learn through this experiment to expand the selection of available VR training modules for multiple logistically and ecologically critical urban sites, including maritime ports.

Materials and Methods

Virtual Environment

We built a custom simulation using the Unity game engine that immerses the user into an underground parking garage with fires blazing on one side and water pouring in from the cracking pipes and slowly flooding the second half of the structure. Fog particles were used to moderately simulate the effect of white smoke while a siren sound is lingering at a range of 60-70 dB and 440Hz-18KHz. Users were tasked to first locate an ambulance before searching for the emergency exit portal. To control for the impact of wayfinding affordances, we created three clones of the simulation. The first two clones featured directional LED lights on either the floor or the walls of the structure, while the third clone did not include any directional lights (Figure 1). Half of the test users played the simulation with the LED lights on (EG experimental group), while the other half was used as a baseline and they played the simulation that did not have any LED lights (CG control group).

Data

We captured data from different modalities using the following three methods. All the raw anonymized data collected for this experiment along with their evaluation metrics and statistical tests are available in GitHub (XXX, 2020).

Behavioral indicators. In-game analytics were collected for all users using the Unity scripting API. The collected features have been widely reported as being useful in prior wayfinding studies. The following user performance metrics are calculated, scripted in the code as game variables and can be readily exported in a structured format for further analysis of statistical significance.

- Completion time: Uses the *Time.time* function to measure the numeric value of the seconds elapsed since the beginning of the simulation.
- Distance travelled: Uses the *Vector3.Distance* to get player movement each frame. These are combined to get the aggregated Euclidian distance. *Vector3* stores a player's position vector (*x,y,z*).
- Average speed: Calculated after the scene is completed simply by the ratio *distance travelled*/*completion time*.
- Total stops: Checked using player position. If the previous frame position is the same as current frame position it is calculated as a stop.
- Number of rotations: Same principle as in total stops. The player orientation is checked each frame and compared to the previous orientation. If the frames have different orientation values it is calculated as one 45 degree rotation. The player rotation control stick turns the

player 45 degrees each time it is pressed. Rotation is obtained from the player object using the function/method "transform.eulerAngles".

- Total rotation: Simple calculation to convert the number of rotations into degrees (one rotation = 45 degrees).
- Ambulance found (Goal 1): Just a flag that is triggered when player reaches the ambulance. The flag also triggers a text prompt to let the user know that the first objective was completed.
- Ambulance time: Time taken until the ambulance is found. It uses the *Time.time* function/method when the player reaches the ambulance trigger.
- Exit found (Goal 2): Same as completion time.

Physiological indicators. Heart rate signal was recorded from 17 subjects throughout the entire experiment using a standard Polar A370 reflective photoplethysmography (PPG) sensor (Figure 3). Polar HR sensors have been used in several studies and have been shown to produce consistent results (Müller et al.,2019; Mercer et al.,2016; Bau et al., 2016; Brooke et al., 2017; Fokkema et al., 2017; Hernández-Vicente et al., 2016; Giles et al., 2016). PPG sensors use light-based technology to read blood volume changes, caused by heart's pumping activity, from tissue capillaries. The signal coming from the sensor gauges the aggregated systolic and diastolic heart activity and has been used extensively in the literature as a key physiological indicator related to the vascular system (Henriksen et al. 2018; Stahl et al. 2016; Jo et al. 2016; Wang et al. 2017; Terbizan, 2002). While this approach lacks the granularity of other beat-to-beat measurements, we can still exploit PPG trend, amplitude and variability, and in conjunction with external stimuli, isolate the impact of potential stress events. Past research in VR games and systems has demonstrated that VR content can induce blood flow velocity and affect PPG amplitude (Rey et al. 2011). Consequently, data from the PPG sensor enables us to explore the role of the signage system as physiological stressor under extreme conditions.

Psychometric indicators. A pre-study questionnaire was used to record the demographics of the subjects and characterize their habits, especially with reference to their acquaintance with games, VR, emergency training, driving, and underground parking facilities. The weight and height of the subjects was also registered for calculating their respective Body Mass Index (BMI). Furthermore, a post-study psychometric game experience questionnaire in English was used to solicit participant's assessment and experience during and after the game in a five-point Likert scale (IJsselsteijn, De Kort, & Poels 2013; Poels et al. 2007).

Participants

Twenty-four subjects were selected to participate in a 20-minute experiment. All users were students and at least 18 years old with no prior experience in VR as confirmed by a self-reported questionnaire. Subjects were selected from a non-probabilistic sample of passersby at the Tellus Arena located in the Linnanmaa Campus of the University of Oulu (Figure 2). This location has been used in numerous VR field trials (Kostakos et al. 2019; Alavesa et al. 2018; Florea et al 2019; Pakanen et al. 2018). Furthermore, we collected HR and behavioral data from a subset of seventeen participants, and psychometric indicators from all twenty-four subjects. The sample from which HR and behavioral data were collected includes non-expert users with no prior VR experience (seldom/never use n=15; monthly use n=2) and some typical gamers (seldom/never $n = 10$; daily/weekly n=7). Furthermore, all participants had a driving license and about half reported using/visiting underground parking facilities. Finally, five males reported to have received military training.

Experimental protocol

The experiments lasted 20 minutes on average and the participants received a coffee voucher worth 2 euros after completing the experiment. After signing a consent form, the Polar A370 HR monitor was fitted with a wrist band around the left arm of the participant and the resting heart rate was recorded. Following Tucker *et al.* 2018, the sensor was applied by the subjects themselves to minimize experimenter contact. Next, the participants were briefed about the nature of the experiment and the content of the game. Subsequently, participants played a short tutorial to learn how to use the controllers and the HMD. After completing the tutorial, participants were assigned to one of the three experimental conditions (no lights, floor lights, and wall lights) (Figure 1) and played the simulation while receiving only basic guidance from the researcher. Subsequent to the conclusion of the game, participants were asked to fill out a psychometric game experience questionnaire. The HR sensor was removed after the last step. HR data was recoded through the experiment to avoid possible artefacts and false positives in the data. For example, previous studies have shown that introducing tasks (e.g. reading) before the actual simulation, results in an upwards trending HR for both the baseline and the treatment groups (Meng and Zhang, 2014; Tucker et al. 2018; Chittaro and Buttussi, 2015).

Working Hypotheses

Using the above sources of data, we use hypothesis testing methods (Welch two Sample t-test, Fisher's exact test, Kruskal-Wallis, and Mann-Kendall) to help us better understand patterns in our data. Specifically, our analysis is driven by the working hypothesis that there will be a statistically significant difference between the performance of the two groups due to the signage system deployed in the experimental version of the game. The hypothesis is largely based on intuition gained from prior statistically significant findings indicating that visual affordances improves the performance of players who are asked to pursue navigation problems in virtual environments. We do not expect to see any differences in the HR data because unlike other studies, our two groups are exposed to the same audiovisual content, thus, any statistically significant difference will only be tied up to the presence or absence of the wayfinding system. These assumptions come along with many limitations that could be addressed and improved in a more rigorous laboratory setting.

Results

Welch's t-test and Student's t-test were carried out to compute the effect of the difference between the means of the experimental group (EG n= 7) and the control group (CG n=10) using the *stats* library in R. Fisher's F-test was used for controlling heteroscedasticity. Continuous demographic, profile, and physiological data for each group are presented in Table 1. There was no significant true difference in means in age and BMI. Looking at the mean HR captured before, during and after the gaming session, participants in the control group had significantly higher rates (Table 1). A Fisher's exact test also revealed no difference in the gender distribution assigned to the two groups (odds=1.31, p=<0.999). Similarly, no differences are noted in the distribution of military training received by subjects across the two groups (odds=0.93, p=<0.999), on the use of underground parking (odds=1.31, p=<0.999), frequency of gaming (odds=0.89, p=<0.646), and familiarity with Virtual Reality (odds=0.68, p=<0.999).

User’s task performance variables captured by the in-game analytics software are presented in Table 2 and Figure 4. A Welch’s t-test shows that there was no statistically significant difference in means when looking at the performance between users who played the simulation with lights (EG) and those who played the simulation with no lights (CG). The lowest *p* value (0.166) was found for the number of turns (Figure 4E), but this result was non-significant. Nevertheless, as shown in Table 2 subjects in the baseline condition made more pauses and turns, walked more and travelled faster, and took more time to find the exit.

TABLE I. CONTINUES DEMOGRAPHIC AND PHYSIOLOGICAL VARIABLES

Group	**CG (n =10)**	**EG (n =7)**	***t* (df)**	***p***	**95% CI**
Age	27.2	26.1	0.39(15)	0.696	-4.608; 6.723
BMI	23	24.8	-0.99(15)	0.334	-5.794; 2.100
HR	88.4	76.7	-63.892 (20888)	<0.0001	-12.11582; -11.39457

TABLE II. GAME PERFORMANCE ANALYTICS VARARAIBLES

Group	**CG (n =10)**	**EG (n =7)**	***t* (df)**	***p***	**95% CI**
Completion time	206	168	0.55(15)	0.587	-108.0815; 183.9958
Walked distance	571	483	0.62(15)	0.541	-212.9697; 389.5697
Average speed	325	300	0.63(11)	0.535	325.5000; 300.7143
Number of stops	53	25	1.14(11.6)	0.274	-24.82213; 79.70785
Number of turns	297	163	1.4(15)	0.166	-62.27264; 330.61550
Time to Goal	24	28	-0.5(14.9)	0.591	-22.34898; 13.20612

TABLE III. PHYSIOLOGICAL TRENDS

Group	***z***	***n***	***p***	***S***	***tau***
Experimental Group EG (n=7)	-13.094	1183	>0.000	-177559	-0.258
Control Group CG (n =10)	43.08	2066	>0.000	1348362	0.6402

The complete time series of the physiological variables related to the HR sensor data is visualized in Figure 5. Preliminary visual inspection of the plots shows that users in the experimental group tend to have a negative trend in their HR while playing the simulation (area marked between the red checkpoints), while users in the control group were observed to have an upwards trend. A Mann-Kendall test was applied to examine the statistical significance of the monotonic trends in HR data between the two both groups (Trend, 2020). The cumulative series are presented in Figure 6 and Table 3. As shown, significant decreasing trends were detected at the 5% significance level for users in the experiment group (z = -13.094), while increasing trends were detected in the HR series for the control group (z = 43.08).

A psychometric game experience questionnaire was issued to a larger sample of 24 users, 11 (6 female, Age m=27.0; SD=6.4) in the control group and 12 (8 female, Age m=28.3; SD=9.7) in the experimental group (IJsselsteijn, De Kort, & Poels 2013; Poels et al. 2007). The questionnaire consists of a core module for in-game (IG) experience and a postgame (PG) module. Kruskal-Wallis test (Table 4) in the postgame module showed statistically significant (p<0.001) difference in negative affect, mean rank being 16.91 in the control group and 30.92 in the experimental group. This suggests a slight negative affect once the simulation was over in the experimental group. In addition, Tension (p=0.001) was 17.70 in the control group and 30.25 in the experimental group.

TABLE IV. IG AND PG GEQ RESULTS

Questionnaire item	**Group**	***n****	**Mean Rank**	***p***
IG Positive affect	EG	55	67.61	0.007
	CG	62	51.36	
	Total	117		
IG Negative affect	EG	44	36.23	<0.001
	CG	50	57.42	
	Total	94		
IG Tension/ annoyance	EG	33	26.92	<0.001
	CG	39	44.60	
	Total	72		
IG Challenge	EG	55	50.07	0.006
	CG	62	66.92	
	Total	117		
PG Negative affect	EG	22	16.91	<0.001
	CG	26	30.92	
	Total	48		
PG Tension	EG	22	17.70	0.001
	CG	26	30.25	
	Total	48		

*n refers to number of answers in each questionnaire item

In the core module measuring in-game experience, the results also showed statistically significant difference in positive affect (p=0.007) and challenge (p=0.006). Mean ranks for positive affect in the control group were 67.61 and for experimental group 51.36. Mean ranks for challenge were 50.07 in the control group and 66.92 in the experimental group respectively. These results indicate that in the experimental group, following the LED lights was possibly perceived as an extra challenge. However, we cannot compare this with how well the users cleared the simulation as there were no statistically significant differences between the aforementioned metrics for this. For in-game experience "negative affect" and "tension and annoyance" p value was less than 0.001. These items are known to correlate in game experience questionnaire (Poels et al. 2007), therefore similar results are not unexpected. Mean rank "tension and annoyance" was 26.92 in the control group and 44.60 in the experimental group. For negative affect mean ranks were 36.23 in the control group and 57.42 in the experimental group. These results suggest that users experienced the experimental condition more negatively.

Discussion

Immersive virtual reality has enabled new opportunities for realistic and inexpensive simulations that can be used for training employees and front-line practitioners in wayfinding, rescues and navigation tasks. The objective of the study was to compare the performance of participants who are immersed into a CBRNE scenario while controlling for experimental condition by means of a signage system. Two groups of participants played the same game with the only difference being the presence/absence of LED lights guiding the users to the emergency exit door. Our initial hypothesis was that users in the experimental group who had the extra help from the LED lights will outperform those from the control group (no lights). Moreover, we hypothesized that the HR data analysis will not return a significant difference between the groups because we did not have strong prior knowledge correlating heart activity with wayfinding affordances. To further quantify our hypotheses, we used evaluation metrics from user's in-game performance, self-reported psychometric questionnaires and physiological readings from a heart rate sensor.

Effect on group selection

Standardized mean difference tests were used to compute the effect of demographic attributes between the baseline and experimental groups where no significant differences were found. This was motivated by prior research observing significant age and gender-related differences in spatial navigation and place learning within virtual environments. Head and Isom (2010) used MRI scans to determine the neural correlates of spatial navigation in a virtual maze in ninety-seven older and young adults. They determined that older adults travelled a greater distance in the virtual environment to locate specific landmarks and acquired less environmental knowledge than the young adults group. In addition, the same study reported significant results in terms of gender deficit with females being able to recall more landmarks than males. In another study with 48 undergraduate and graduate students, Sandstrom et. al (1998) observe that females perform better in navigation tasks when landmark information was presents, but unlike males, their performance "was severely disturbed" by the absence of landmark or the introduction of unreliable landmark information. Similarly, Kober and Neuper (2011) analyzed EEG data from 27 subjects immersed into a VR maze and identified differences in cortical theta activity between the two gender groups with female participants generating stronger increase in theta oscillations when using landmarks as navigation aids. Additionally, motivated by Head and Isom (2010) we did not find significant differences between the means for other variables like gaming time, familiarity with VR technology, driving skills, car ownership, underground parking frequency or body mass index (BMI).

Nevertheless, t-test showed a significant variation in HR values, with participants in the baseline conditions having larger means. Various factors can be attributed to this variability including outliers that derive from motion sickness felt by users, subjects with high BMI known to have high resting heart rate, and participants who engage in intense athletic activities who are known to have low resting heart rate.

Effect on behavioural outcomes

Prior research has found strong correlations between wayfinding performance and visual sensory cues in VR installations (Radianti et al, 2020; Feng et al., 2019; Williams-Bell et al. 2015). In the following experiment, Tang et al. (2009) created a game with an escape goal to evaluate the impact of green-lighted exit signs on evacuation time. Their study included three indoor scenarios (no sign, old sign, new sign) and was tested with one hundred and seven subjects. The presence of either of the two exit sign types, according to their results, reduced the latency of players to finding the exit portal. In another experiment, Vilar et al. (2014) recruited fifty-four volunteers to test a VR wayfinding apparatus that included a dual signage system with color lights on the floor and on the wall of a building. Their experiment captured in-game user performance for distance travelled, time spent, pauses and average speed. Like many of the previous studies, they report statistically significant differences for distance travelled and time spent but reported mixed results for average speed and pauses. In a completely different setting, Cosma et al. (2016) measured the mobility paths of participants to analyze evacuation behavior using stripes of dynamic lights in an underground rail tunnel. Statistical tests disclosed that the wayfinding path of users was only marginally affected by these installations and found no differences for total evacuation and interaction times.

Our findings indicate that the performance and behavior of participants as captured by the in-game software is not statistically significant, suggesting that the presence/absence of assistive LED lights had no significant direct effect in the completion of goals. Although the deficit found is not statistically significant, our findings also point out that subjects assigned to the control condition spent more time finding the exit, made more pauses and walked more and faster than the group assigned to the experimental condition. Overall, our results are more aligned with experiments conducted in underground locations and less in tune with those simulations run inside buildings. The theory of affordance, as already suggested in prior studies from the field, might partly explain this misalignment. Underground locations like evacuation tunnels, storage tanks, garages and other industrial sites in general, often lack landmark information that can be used as navigational affordance to help users select from a number of potential navigation paths (Bonner and Epstein, 2017).

Effect on Heart Rate outcomes

Results from the Mann-Kendall test show that HR values increase significantly for the control group while dropped for the experimental group. This indicates that although the LED lights might not have significant impact on the in-game performance, they nevertheless, significantly affect users' physiology. This is a novel finding that has not been observed in prior related works. Particularly, in a related experiment, Meng and Zhang (2014) collected HR data from forty participants exposed to a panorama manifestation system simulating a fire evacuation in a virtual hotel environment. The treatment group was exposed to stimuli like fire alarm, virtual fire, smoke and sporadic explosions. The signage system in both control and treatment group remained the same. They observed that the HR for both groups increased once the wayfinding simulation started, with the treatment group having overall significantly higher HR ratio than the control group. The authors discuss this finding in the context of the stressors induced by the presence of multisensory stimuli that, as anticipated, elevated the HR. Similar findings have been observed by Chittaro and Buttussi (2015) who reported significant variations in blood volume pulse amplitude between users who played a VR serious game simulating airplane evacuation procedures and users who just browsed the information on a safety card. Overall, prior experiments make explicit linkages between visual stimuli and HR (Tucker et al., 2018). This link is now been commercially exploited by the gaming industry (World Economic Forum, 2018).

Against this backdrop, our experiment held all gaming conditions constant besides the presence/absence of wayfinding signs. This assertion is also evident from the analysis of self-reported psychometric data showing that the experimental group experienced the game drastically different (more negatively) than the baseline group. By extension, we are confident to assert that the difference in HR data seen in our results can be attributed to a latent cognitive perception mechanism for which prior literature exists. For instance, cognitive workload and heart rate (HR) have been explored in a driving simulation developed in Hidalgo-Muñoz et al. (2019). Subjects were exposed to tasks demanding low cognitive workload (beep counting) and high cognitive workload with mental displacement (combination of memorization and arithmetic) while driving and shifting. ECG signal shows that HR increases when performing the dual activity (i.e when driving). The authors suggest that this can be explained by the competition of working memory components prompted by the multiple cognitive tasks.

Thus, we reason that in our experimental set up, the decreasing heart rate can be explained within the context of a decreasing cognitive workload prompted by the presence of the wayfinding signage, enabling the subjects to form faster spatial representations during the wayfinding task. Prior works that tackle brain connectivity through the use of EEG data provide a more fine-grained view into the mechanism that juxtaposes cognitive workload, wayfinding affordances and physiological variables. For instance, Sharma *et al.* (2017) studied the occurrence of errors in a VR wayfinding simulation with landmarks. After analyzing EEG data, the authors concluded that the experimental group exposed to the simulation with landmarks had better brain network efficiency and higher left-hemispheric activation with higher activity in the theta band known to be linked with landmark processing, object recognition and categorization. Contrarily, the control group (no landmarks) exhibited a higher number of active electrodes in the left posterior inferior and superior regions of the brain, with the authors reasoning that the subjects in this group had a higher chance of disorientation - which increased the re-wiring cost associated with such frequent cognitive demands.

CONCLUSION

The paper presents the results of a VR experiment with twenty-four participants, simulating an underground parking wayfinding installation in Oculus Rift S HMD. We found that assistive lights have a greater effect on physiological outcomes (heart rate) than they do on behavioral ones (game performance). One possible explanation is that there was more efficient brain wiring and cognitive processing in the experimental group, where wayfinding affordances were present, and thus, easier for the users to find the exit portal. This is in line with results from the psychometric self-reported data. Participants in the experimental condition also reported to have received a more negative gaming experience. This inconsistency can be attributed to subjects in this condition having less opportunities for exploring the scene, lack of a novelty effect or not meeting their hyped expectations. This conclusion has implications on how wayfinding tasks (and games in general) are designed and evaluated. Many current studies use self-reported metrics and psychometric questionnaires to asses user experience in VR. The inconsonance we have found between self-reported data and biometric measurements is alarming and warrants re-inspection of many current findings in VR related research.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Partly funded by ##### and ####

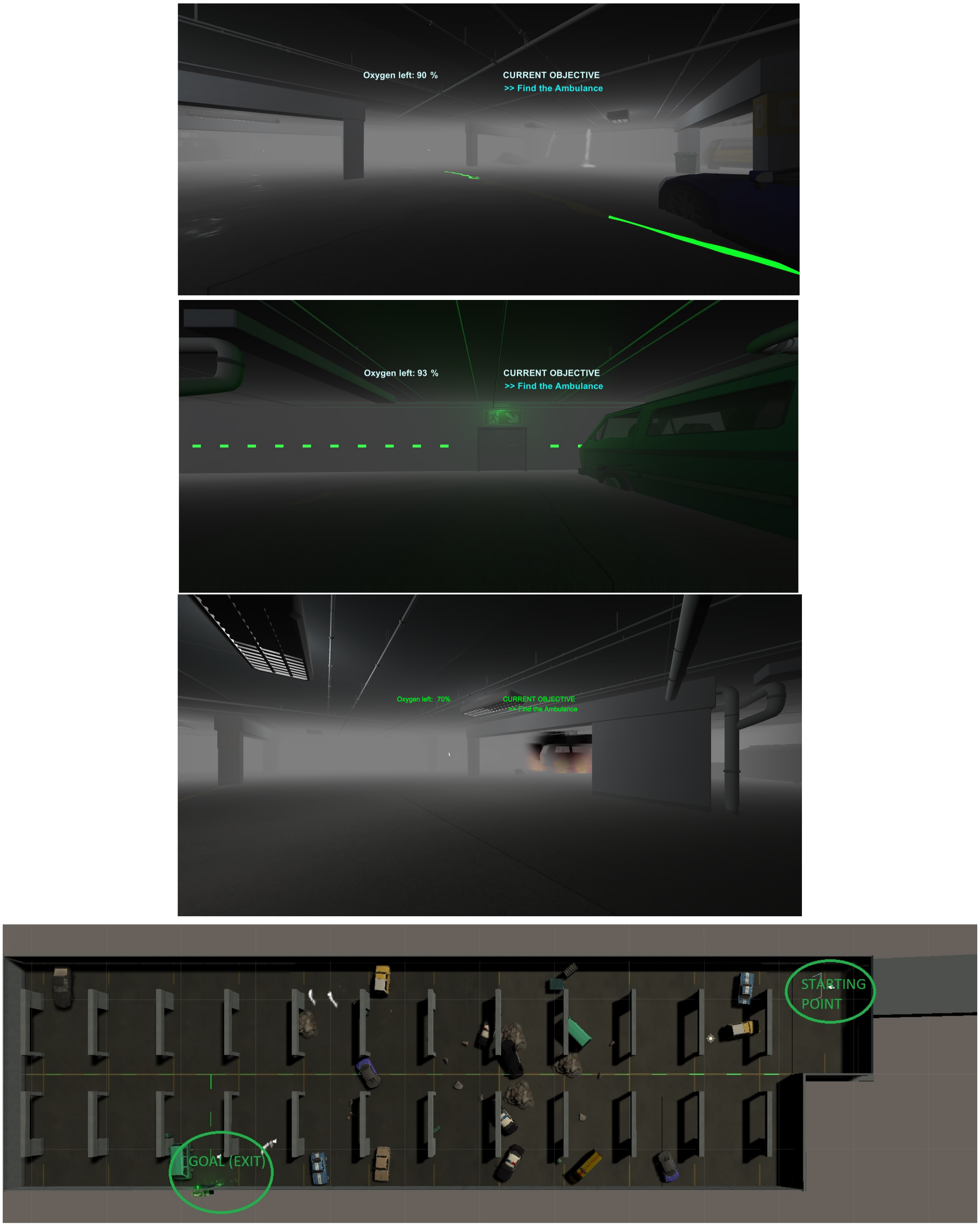

Fig. 1. From top to bottom: Participants' view of the underground parking scenes in VR along with associated game objects and wayfinding system. The virtual environment and game objects are created based on requirements produced by the end-users of the XXX project.

Fig. 2. Ubiquitous space (Cube) in Tullus repurposed for the requiremnts of the VR experiemnts. The Cube is sound proof and has autonomous ventilation and temperature controls.

Fig. 3. Polar A370 fitness tracker featuring a reflective PPG sensor consistsing of a light source and a photodetector that can sense blood volume changes from tissue capillaries.

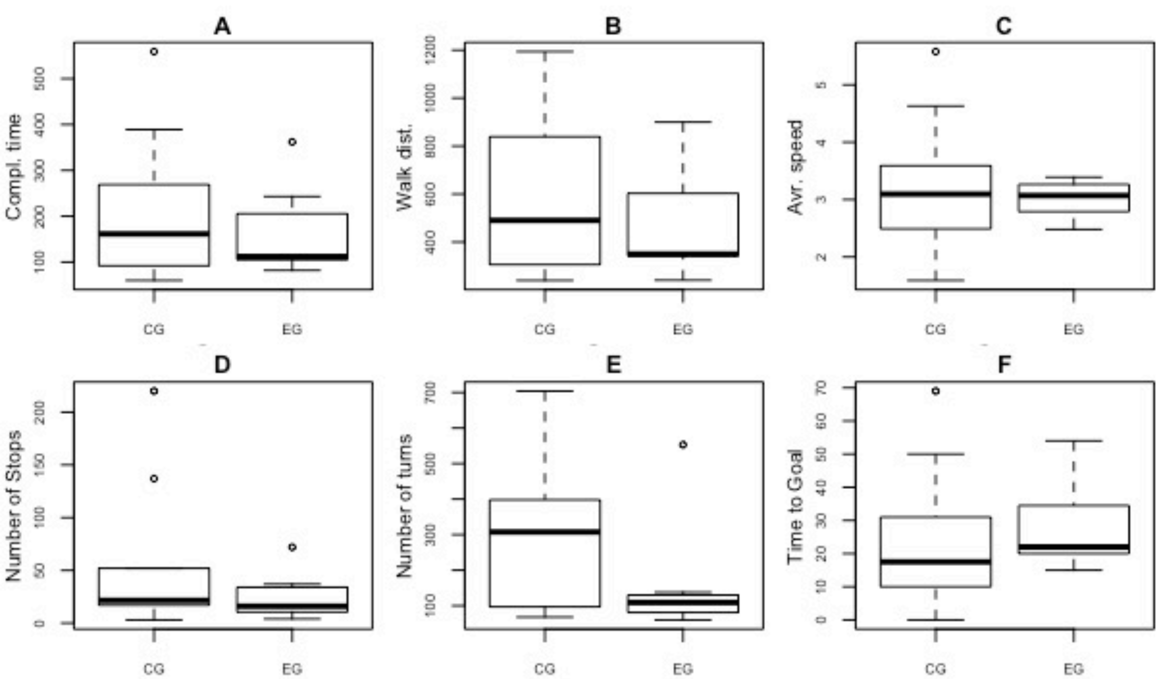

Fig. 4. Boxplots with means for in-game analytics variables. The following variables are shown for the Experimental (EG) and Control (CG) groups: absolute time requited by each user complete the simulation (A), Absolute distance traveled by user during the simulation (B), Average speed of user measured as m/sec. (C), number of stops (D), number of turns (E), and time in seconds required by user to find the ambulance (goal) (F).

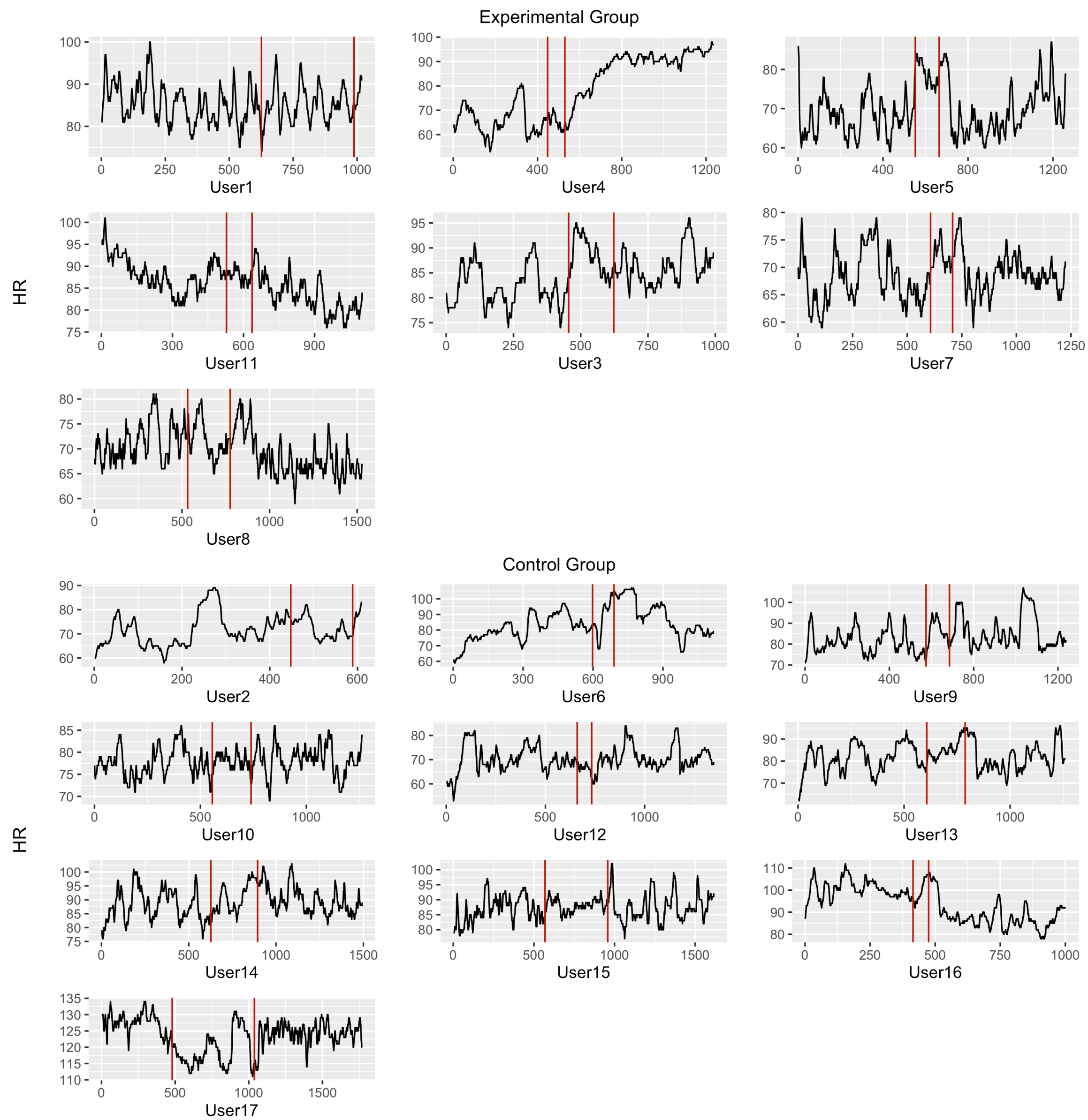

Fig. 5 Raw Heart Rate data as recorded by the Polar sensor. Bottom grid shows the HD data for the 10 particiapnts who tested the control similation of the game. The top grid shows raw HR data for the seven users in the experimental group. The x axis in all graphs shows the time elapsed in seconds. The actual simulation takes place in the intercepts marked with red checkpoints. Note that user 4 suffered from motion seeknes as reported after the end of the simulation.

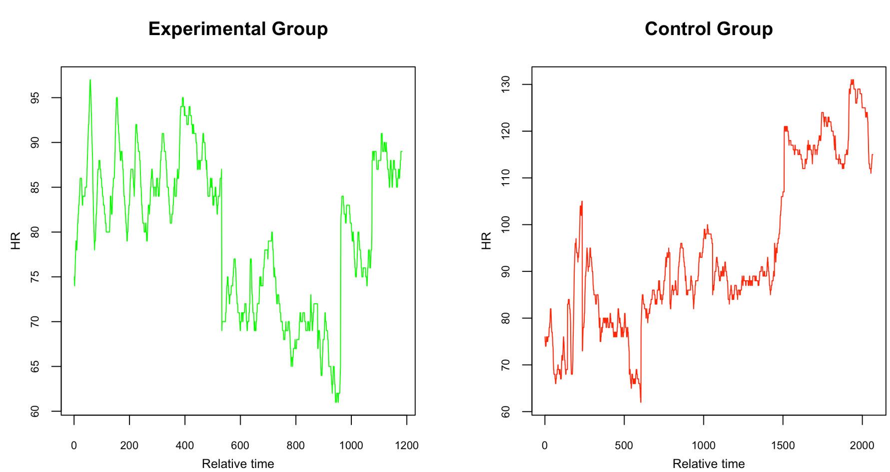

Fig. 6. Cumulative raw HR data as recorded by the Polar sensor during the navigation task.

Bibliography

Alavesa, P., Minna Pakanen, Arttu Niemelä, Weiping Huang, Satu Väinämö, Lotta Haukipuro, Leena Arhippainen, and Timo Ojala. 2018. Mobile Augmented Reality Client as a UX Method for Living Lab's User Involvement Tool. In Proceedings of the 22nd International Academic Mindtrek Conference (Mindtrek '18). ACM, New York, NY, USA, 135-142. DOI: https://doi.org/10.1145/3275116.3275147

Arias, S., La Mendola, S., Wahlqvist, J. et al. Virtual Reality Evacuation Experiments on Way-Finding Systems for the Future Circular Collider. Fire Technol 55, 2319–2340 (2019). https://doi.org/10.1007/s10694-019-00868-y

Bai, Y., Welk GJ, Nam YH, Lee JA, Lee JM, Kim Y, et al. Comparison of consumer and research monitors under semistructured settings. Med Sci Sports Exerc 2016;48(1):151-158.

Bonner MF & Epstein RA (2017) Coding of navigational affordances in the human visual system. Proceedings of the National Academy of Sciences. 114(18): 4793-8. https://doi.org/10.1073/pnas.1618228114

Brooke, SM, An HS, Kang SK, Noble JM, Berg KE, Lee JM. Concurrent validity of wearable activity trackers under free-living conditions. J Strength Cond Res 2017 Apr, 31(4):1097-1106.

Chieh-Hsin Tang, Wu-Tai Wu, Ching-Yuan Lin, Using virtual reality to determine how emergency signs facilitate way-finding, Applied Ergonomics,Volume 40, Issue 4, 2009, Pages 722-730, https://doi.org/10.1016/j.apergo.2008.06.009.

Chittaro L., Buttussi F., Assessing Knowledge Retention of an Immersive Serious Game vs. a Traditional Education Method in Aviation Safety, IEEE Transactions on Visualization and Computer Graphics, vol. 21, no. 4, 2015, pp. 529– 538.

Christopher, K., Port security management, 2nd ed. CRC Press, Taylor Francis, 2014.

Cosma, G., Enrico Ronchi & Daniel Nilsson (2016) Way-finding lighting systems for rail tunnel evacuation: A virtual reality experiment with Oculus Rift®, Journal of Transportation Safety & Security, 8:101-117, DOI: 10.1080/19439962.2015.1046621

Donna J. Terbizan, Brett A. Dolezal & Christian Albano (2002) Validity of Seven Commercially Available Heart Rate Monitors, Measurement in Physical Education and Exercise Science, 6:4, 243-247, DOI: 10.1207/S15327841MPEE0604_3

Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, Hamilton CB, Li LC Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data JMIR Mhealth Uhealth 2018;6(8):e10527 DOI: 10.2196/10527

Feng, Z. Vicente A. González, Robert Amor, Ruggiero Lovreglio, Guillermo Cabrera-Guerrero, Immersive virtual reality serious games for evacuation training and research: A systematic literature review, Computers & Education, Volume 127, 2018, Pages 252-266, https://doi.org/10.1016/j.compedu.2018.09.002.

Feng, Zhenan, Vicente A. González, Robert Amor, Ruggiero Lovreglio, and Guillermo Cabrera-Guerrero. 2018. “Immersive Virtual Reality Serious Games for Evacuation Training and Research: A Systematic Literature Review.” Computers & Education 127:252–66.

Florea, C. et al., "Extending a User Involvement Tool with Virtual and Augmented Reality," 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 2019, pp. 925-926.

Fokkema, T., Kooiman TJ, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers depend on walking speed. Med Sci Sports Exerc 2017 Apr, 49(4):793-800.

Giles D, Draper N, Neil W. Validity of the Polar V800 heart rate monitor to measure RR intervals at rest. Eur J Appl Physiol 2016 Mar;116(3):563-571

Head, D., Marlisa Isom, Age effects on wayfinding and route learning skills, Behavioural Brain Research, Volume 209, Issue 1, 2010, Pages 49-58, https://doi.org/10.1016/j.bbr.2010.01.012.

Henriksen A, Haugen Mikalsen M, Woldaregay AZ, Muzny M, Hartvigsen G, Hopstock LA, Grimsgaard S, Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables, J Med Internet Res 2018;20(3):e110, DOI: 10.2196/jmir.9157

Hernández-Vicente, A., Santos-Lozano A, De Cocker K, Garatachea N. Validation study of Polar V800 accelerometer. Ann Transl Med 2016 Aug;4(15):278

Hidalgo-Muñoz Antonio R., Béquet Adolphe J., Astier-Juvenon Mathis, Pépin Guillaume, Fort Alexandra, Jallais Christophe, Tattegrain Hélène, Gabaude Catherine, Respiration and Heart Rate Modulation Due to Competing Cognitive Tasks While Driving, Frontiers in Human Neuroscience, 12, 2019, DOI: 10.3389/fnhum.2018.00525

IJsselsteijn, W. A., Y. A. W. De Kort, and Karolien Poels. 2013. "*The Game Experience Questionnaire*." Eindhoven: Technische Universiteit Eindhoven 3–9.

Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of biofeedback wearables for photoplethysmographic heart rate tracking. J Sports Sci Med. 2016 Sep;15(3):540–547.

Kober, S.E., Christa Neuper, Sex differences in human EEG theta oscillations during spatial navigation in virtual reality, International Journal of Psychophysiology, Volume 79, Issue 3, 2011, Pages 347-355, https://doi.org/10.1016/j.ijpsycho.2010.12.002.

Kostakos, P., Paula Alavesa, Jonas Oppenlaender, and Simo Hosio. 2019. VR ethnography: a pilot study on the use of virtual reality "go-along" interviews in Google street view. In Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia (MUM '19). Association for Computing Machinery, New York, NY, USA, Article 53, 1–5. DOI:https://doi.org/10.1145/3365610.3368422

Lynch, K. 1960. The Image of the City. Cambridge, MA: MIT/Harvard University press.

May M. (2004) Wayfinding, Ships and Augmented Reality. In: Andersen P., Qvortrup L. (eds) Virtual Applications. Springer, London, https://doi.org/10.1007/978-1-4471-3746-7_10

Meng, F. & Wei Zhang (2014) Way-finding during a fire emergency: an experimental study in a virtual environment, Ergonomics, 57:6, 816-827, DOI: 10.1080/00140139.2014.904006

Meng, F. & Wei Zhang (2014) Way-finding during a fire emergency: an experimental study in a virtual environment, Ergonomics, 57:6, 816-827, DOI: 10.1080/00140139.2014.904006

Mercer, K., Li M, Giangregorio L, Burns C, Grindrod K. Behavior change techniques present in wearable activity trackers: a critical analysis. JMIR Mhealth Uhealth 2016;4(2):e40

Müller, A. M., Wang, N. X., Yao, J., Tan, C. S., Low, I., Lim, N., Tan, J., Tan, A., & Müller-Riemenschneider, F. (2019). Heart Rate Measures From Wrist-Worn Activity Trackers in a

Laboratory and Free-Living Setting: Validation Study. JMIR mHealth and uHealth, 7(10), e14120. https://doi.org/10.2196/14120

Pakanen, M., Lovén, L., Alavesa, P., Gilman, E., Terken, J., Eggen, B., & Pirttikangas, S. (2018). Design challenges of wellbeing supporting smart environment in collaborative use situations. In UbiComp/ISWC 2018 - Adjunct Proceedings of the 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2018 ACM International Symposium on Wearable Computers (pp. 688-692). Association for Computing Machinery, Inc. https://doi.org/10.1145/3267305.3267691

Poels, K., de Kort, Y. A., & IJsselsteijn, W. A. (2007). D3. 3: Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games.

Radianti, J., Tim A. Majchrzak, Jennifer Fromm, Isabell Wohlgenannt, A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda, Computers & Education, Volume 147, 2020, 103778, ISSN 0360-1315, https://doi.org/10.1016/j.compedu.2019.103778.

Rey, B., Parkhutik, V., and Alcañiz, M. (2011). Breaks in presence in virtual environments: an analysis of blood flow velocity responses. *Presence* 20, 273–286. doi: 10.1162/PRES_a_00049.

Rogers, S., “2019: The Year Virtual Reality Gets Real,” Forbes. [Online]. Available: https://www.forbes.com/sites/solrogers/2019/06/21/2019-the-year-virtual-reality-gets-real/.

Ronchi, E, Mayorga, D, Lovreglio, R, Wahlqvist, J, Nilsson, D. Mobile-powered head-mounted displays versus cave automatic virtual environment experiments for evacuation research. Comp Anim Virtual Worlds. 2019; 30:e1873.

Ronchi, E., Nilsson, D., Kojić, S. et al. A Virtual Reality Experiment on Flashing Lights at Emergency Exit Portals for Road Tunnel Evacuation. Fire Technol 52, 623–647 (2016). https://doi.org/10.1007/s10694-015-0462-5

Ruddle, R. A. and S. Lessels, "Three Levels of Metric for Evaluating Wayfinding," in Presence, vol. 15, no. 6, pp. 637-654, 1 Dec. 2006.

Sandstrom, N.J , Jordy Kaufman, Scott A. Huettel, Males and females use different distal cues in a virtual environment navigation task, Cognitive Brain Research, Volume 6, Issue 4,1998, Pages 351-360, https://doi.org/10.1016/S0926-6410(98)00002-0.

Sharma, G., Kaushal Yash, Chandra Sushil, Singh Vijander, Mittal Alok P., Dutt Varun, Influence of Landmarks on Wayfinding and Brain Connectivity in Immersive Virtual Reality Environment, Frontiers in Psychology, 8, 2017, DOI:10.3389/fpsyg.2017.01220.

Stahl SE, An HS, Dinkel DM, Noble JM, Lee JM. How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? BMJ Open Sport Exerc Med. 2016;2(1):1–7. doi: 10.1136/bmjsem-2015-000106.

Su, F. and I. Anthony, “Reassessing CBRN Threats in a Changing Global Environment”. Solna, Sweden: Stockholm International Peace Research Institute, 2019. https://www.sipri.org/sites/default/files/2019-06/1906_cbrn_threats_su_anthony_0.pdf

Tang, C.-H., W.-T. Wu, and C.-Y. Lin, “Using virtual reality to determine how emergency signs facilitate way-finding,” Applied ergonomics, vol. 40, no. 4, pp. 722–730, 2009.

Trend (2020) R package ‘trend’, January 13, 2020, https://cran.r-project.org/web/packages/trend/trend.pdf

Tucker, K.L. Marsh, T. Gifford, X. Lu, P.B. Luh, R.S. Astur, The effects of information and hazard on evacuee behavior in virtual reality, Fire Safety Journal, Volume 99, 2018, Pages 1-11, ISSN 0379-7112, https://doi.org/10.1016/j.firesaf.2018.04.011.

van Buuren, S. and D. J. Wijnmalen, "Measuring Psychosocial Impact of CBRN Incidents by the Rasch Model," Journal of applied measurement, vol. 16, p. 3, 2015.

Vilar, E., Rebelo, F. and Noriega, P. (2014), Indoor Human Wayfinding Performance Using Vertical and Horizontal Signage in Virtual Reality. Hum. Factors Man., 24: 601-615. doi:10.1002/hfm.20503

Wang R, Blackburn G, Desai M, et al. Accuracy of Wrist-Worn Heart Rate Monitors. JAMA Cardiol. 2017;2(1):104–106. doi:10.1001/jamacardio.2016.3340

Williams-Bell, F.M., Kapralos, B., Hogue, A. et al. Using Serious Games and Virtual Simulation for Training in the Fire Service: A Review. Fire Technol 51, 553–584 (2015). https://doi.org/10.1007/s10694-014-0398-1

World Economic Forum (2018) This VR horror game is only as scary as your heart can handle https://www.weforum.org/agenda/2018/08/this-vr-horror-game-is-exactly-as-scary-as-your-body-can-handle

Zhao, CM, Siu Ming Lo, SP Zhang, and M. Liu. 2009. "A Post-Fire Survey on the Pre-Evacuation Human Behavior." Fire Technology 45(1):