

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

Panos Kostakos, Paula Alavesa, Mikko Korhikoski (University of Oulu, Finland)

Mario Monteiro Marques, Victor Lobo (CINAV-Portuguese Navy Research Center, Portugal)

Filipe Duarte (Optimal Defence, Portugal)

### **Abstract**

*Background.* **Wayfinding** has been adopted in several intense **evacuation and navigation simulations**; however, the use of biometric measurements for characterizing physiological outcomes has been rather overlooked and applied under limited laboratory conditions.

*Methods.* 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 control group was exposed to a scene without assistive lights, and the experimental condition group was exposed to the same scene with assistive lights.

*Results.* 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 control group displaying an increasing HR trend during simulation.

*Conclusions:* 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:** Wayfinding, affordances, assistive lights, evacuation and navigation simulations, underground parking, heart rate (HR), Polar A370

## Introduction

Emergencies are critical incidents with a high potential for prolonged negative social, economic, and political impacts (Feng, González, Amor, Lovreglio, & Cabrera-Guerrero, 2018). Chemical, Biological, Radiological, Nuclear, and high-yield Explosive (CBRNE) events are extreme emergencies that can destabilize governments, lead to conditions that aggravate violence, or promote terrorism. Furthermore, these events can quickly wear down the civil protection infrastructure and the capability of the first responders—leading to a loss of trust in public institutions (van Buuren & Wijnmalen, 2015; Su & Anthony, 2019). Ports are a vital 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). Immersive virtual reality (VR) provides a way to simulate emergencies, including CBRNE events, without causing physical harm.

Simulations and serious games allow people to train on realistic emergencies and help to improve their chances of staying operational in high-risk situations. While simulations for training and therapy are the most widely researched application areas of virtual reality (VR), there yet remain many unanswered questions. The emotional impact of simulations and the resulting influence on performance is one of the less explored topics in the literature, possibly because of

the complexity of the required research setups and because of the necessity for mixed-method approaches combining different complementing measurements for comprehensive results.

The research presented in this article builds on the XXXX project (XXX, XXX), a project which aims to support front-line emergency medical responders and law enforcement/security services and ship crews by providing them with virtual reality simulations (VRS). VRSs increase the immersive experience of training scenarios through the use of realistic physiological and emotional responses. Therefore, many public and private organizations across the globe are adapting VR systems to train responders on dangerous situations that cannot be illustrated accurately using traditional media. VR training, featuring evacuation and wayfinding scenarios, can be carried out in realistic environments, such as a seaport or other sites under construction without disrupting the workflow in the organization (Tang, Wu, & Lin, 2009). For example, the European Organization for Nuclear Research recently conducted a feasibility study for the construction of the Future Circular Collider (FCC); it included VR solutions for evacuation wayfinding experiments in large underground tunnels (Arias et al., 2019).

In this study, we designed a wayfinding experiment in virtual reality using the Oculus Rift S head-mounted display (HMD) (Oculus, 2020). Twenty-four participants were presented an immersive VR simulation of a burning underground parking lot and were asked to navigate to the exit. Participants were split into two groups. The control group was exposed to a scene without assistive lights, while the experimental condition group was exposed to the same scene with assistive lights. The assistive lights were mounted on either floor or the walls of the garage, pointing towards the emergency exit. We measured behavioral, physiological, and psychological outcomes to investigate the effects of wayfinding assistive lights. 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 harbor areas, or in particular, aboard ships, cruise vessels, and ferries. Our main goal is to use what we have learned from this experiment to expand the selection of available VR training modules for multiple logistically and ecologically critical urban sites, including maritime ports.

## **Related work**

### *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. Lynch (1960) defined *wayfinding* as “the consistent use and organization of definite sensory cues from the external environment” (p. 3). To put this in everyday terms, wayfinding is the conscious use of architectural elements for navigation and orientation in space and time. Furthermore, because architectural elements have the property of being subject to individual interpretation, wayfinding also falls within the scope of affordance theory (Gibson, 1977; 2015; Norman, 2013), which describes how shapes and objects can prompt human actions. Consequently, the lexical compound *wayfinding affordances* has a double meaning: it describes the literal use of cues for space/place orientation, and it measures the usability of those cues to promote optimal actions and behavioral responses (Dondlinger & Lunce, 2009; Gibson, 1977).

There is a vast literature about 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 have spurred interest in navigation experiments using immersive virtual reality. The emerging literature describes VR systems that contain contraptions ranging from mobile, desktop, and large displays to highly immersive HMDs, Fish Tank and Cave Automatic Virtual Environment (CAVE), and panorama manifestation (PM) systems (Lovreglio, Wahlqvist, & Nilsson, 2019; Meng & Zhang, 2014; Ronchi et al., 2016; Ronchi, Mayorga).

### *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 (Feng et al., 2019; Fromm, & Wohlgenannt, 2020; Radianti, Majchrzak, Williams-Bell et al., 2015); and (3) self-reported psychometric and reasoning measures, archived with several qualitative protocols including go-along (Feng et al., 2018; Kostakos, Alaves, Oppenlaender, & Hosio, 2019) and think-aloud interviews and structured questionnaires (Ruddle & Lessels 2006). In regard 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: (1) distance travelled by the user in order to complete the objective/mission of the game, (2) time required to complete the objective/mission, (3) pauses during the simulation that are calculated as user inactivity in seconds or frames, (4) speed of the user, measured as the total distance travelled divided by time spent, and (5) total area covered by the user during the simulation.

Besides behavioral indicators, a few studies have also used biometric measurements for characterizing physiological outcomes that can be beneficial in measuring stress and anxiety levels, although under limited laboratory conditions (Chittaro & Buttussi, 2015; Meng & Zhang, 2014). For instance, recent systematic reviews on evacuation experiments using VR have listed a handful of studies that analyzed the impact of the simulation on cardiac physiological outcomes (Feng et al., 2019; Radianti et al., 2020; Williams-Bell et al., 2015).

Meng and Zhang (2014) measured heart rates of 40 participants immersed into a fire evacuation scenario in a virtual hotel and found a physiological outcome in the form of higher heart rate. 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 (HR) 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 immediately after the start of the simulation—a pattern that seems to be consistent in 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. Similarly, 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 control group who 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 of how best to 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.

This article 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 consisted 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 that exist in the current literature (Feng et al., 2019; Radianti et al., 2020; Williams-Bell et al., 2015), we introduced the use of a photoplethysmogram (PPG) sensor for monitoring changes in HR levels. The widespread availability of wearable and cost-effective biosensors like PPG electroencephalogram (EEG) 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). We chose this scenario 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, which include a pier, a small village, and other buildings.

## **Methods**

## Participants

Twenty-four subjects (n=24) were selected to participate in a 20-minute experiment. All users were students and at least 18 years old, with insignificant 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 (Alavesa et al., 2018; Florea et al., 2019; Kostakos et al., 2019; Pakanen et al., 2018). Data on psychometric indicators were collected from all twenty-four subjects. Due to instrumentation failure, HR and behavioral data were collected from a subset of seventeen participants (n=17). The sample included students from across disciplines with no prior VR experience. Furthermore, all participants had a driving license, and about half reported using/visiting underground parking facilities. Finally, five males reported military training.

## Virtual Environment

We built a custom simulation using the Unity game engine (Unity, 2020) that immerses the user in an underground parking garage with fires blazing on one side and water pouring in from the cracking pipes and slowly flooding the other half of the structure. Fog particles were used to moderately simulate the effect of white smoke, while a siren sound was lingering at a range of 60-70 dB and 440Hz-18KHz. Users were asked to first locate an ambulance (Goal 1) and then find the emergency exit portal (Goal 2). We created three clones of the simulation in order to examine the impact of wayfinding affordances (the independent variable). 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). Subsequently, we tested the simulations using a between-



group study design with two conditions: Half of the test users played the simulations with LED lights (EG experimental group), while the other half (the control group) played the simulation without any LED lights (CG control group). Furthermore, to control for the possible confounding effect of exposure to the signage system previously reported by Vilar, Rebelo, and Noriega (2014), half of the users in the experimental group played the clone with the floor lights, while the other half played the simulation with wall lights. With regard to the overall colourscape of the simulation, the colour of the directional LED lights was high-intensity green (HEX #0DFC1D) placed against a grey backdrop (garage floor or walls). The remaining environment was realistically lighted and coloured, following accessible colour recommendations for VR applications (W3C RQTF, 2020), and the best practices for indoor wayfinding tasks (Hidayetoglu, Yildirim & Akalin, 2012).

## Measures

We captured data from different modalities using the three methods discussed below. 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 application programming interface (API) (Unity API, 2019). The collected features have been widely reported to be useful in prior wayfinding studies (Cosma, Ronchi, & Nilsson, 2016; Vilar et al., 2014). The following user performance metrics are calculated, scripted in the code as game variables, and are readily exported in a structured format for further analysis of statistical significance. Users' avatars were tracked through the Oculus Touch Controllers (Oculus, 2020).

- Completion time: Uses the *Time.time* function to measure the numeric value of the seconds elapsed from the beginning of the simulation.
- Distance travelled: Uses the *Vector3.Distance* to get player movement for each frame. These are combined to get the aggregated Euclidian distance. *Vector3* stores a player's position vector  $(x,y,z)$ .
- Average speed: Is calculated after the scene is completed through the simple ratio of *distance travelled/completion time*.
- Total stops: Is 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: The same principle is applied as in total stops. The player orientation is checked for 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): This is simply a flag that is triggered when a player reaches the ambulance. The flag also triggers a text prompt to let the user know that the first objective has been 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): This is simply a flag that is triggered when a player reaches the exit.

**Physiological indicators.** Heart rate signal was recorded for 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 (Bai et al., 2016; Brooke et al., 2017; Fokkema, Kooiman, Krijnen, van der Schans, & De Groot 2017; Giles, Draper, & Neil, 2016; Hernández-Vicente, Santos-Lozano, De Cocker, & Garatachea, 2016; Mercer, Giangregorio, Burns, & Grindrod, 2016; Müller et al., 2019). 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; this system has been used extensively in the literature as a key physiological indicator for the vascular system (Henriksen et al., 2018; Jo, Lewis, Directo, Kim, & Dolezal, 2016; Stahl, An, Dinkel, Noble, & Lee, 2016; Terbizan, Dolezal, & Albano, 2002; Wang et al., 2017). Furthermore, because "stress hormones" (Adrenaline, Cortisol, and Norepinephrine) are known to set off a chain of events that lead to elevated heart rate, wearable PPG sensors have also been used to capture mental stress under different laboratory conditions, including virtual reality environments (Cho et al., 2017). While this approach lacks the granularity of other beat-to-beat measurements, it is still possible to 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, Parkhutik, & Alcañiz, 2011). Consequently, data from the PPG sensor enables us to explore the role of the signage system as physiological stressor under extreme conditions. The weight and height of the subjects was registered for calculating their respective Body Mass Index (BMI) which is known to have an influence on resting HR and heart rate variability (HRV) (Föhr et al., 2016).

**Psychometric indicators.** A post-study game experience questionnaire (GEQ) (IJsselsteijn, De Kort, & Poels, 2013; Poels, Kort & Ijsselsteijn, 2007) in English was used to solicit participants' assessment and experience during and after the game. GEQ has been widely used in studying user experience in virtual worlds and games (Johnson et al., 2018; Law et al., 2018) and contains four modules: 1) core questionnaire, 2) social presence, 3) in-game module, and 4) post-game module. All modules are to be issued right after the experience. The core and in-game modules require the user to reflect how they felt during the experience, while the post-game module is for the user to reflect how they feel after the experience. We used the core, in-game, and post-game modules in this study. The in-game experience measuring modules contain seven items for "competence, sensory and imaginative immersion, flow, tension/annoyance, challenge, negative affect, and positive affect" (IJsselsteijn et al., 2013, p. 3). The post-game module contains items for positive experience, negative experience, tiredness, and returning to reality. In the results, we only report those items in GEQ, that gave us statistically significant differences between the experimental conditions. In addition, we used a pre-study questionnaire to gather demographic data about the participants' age, gender, height, weight, major/profession, military service, and voluntary rescue work. The questionnaire also elicited socio-demographic information to characterize participants' habits, especially with reference to their acquaintance with games, VR, emergency training, driving, and underground parking facilities.

## Data Analysis Methods

Using the above sources of data, we used 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 was driven by the working hypothesis that there will be a statistically significant difference between the performance of the two groups (lights/no lights) because of the signage system deployed in the experimental version of the game. The hypothesis was largely based on intuition gained from prior statistically significant findings that indicated that visual affordances improved the performance of players who are asked to pursue navigation problems in virtual environments (Feng et al., 2018; Radianti et al., 2020; Tang et al., 2009; Vilar et al., 2014; Williams-Bell et al., 2015). We did not expect to find 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 be the result of the presence or absence of the wayfinding system. In addition to these assumptions, there were many limitations that could be addressed and improved in a more rigorous laboratory setting.

## Procedure

The experiments lasted 20 minutes on average, and the participants received a coffee voucher worth two euros after completing the experiment. In order to maintain a balanced sample, new participants were grouped into the three experimental conditions by means of a tally sheet that contained the gender and experimental group of previous participants. After signing a consent form, the Polar A370 HR (Polar, 2020) monitor was fitted with a wrist band around the left arm of the participant, and the resting HR 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. Then, using the Oculus Rift S controller's joystick for free locomotion inside the game scene, 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. Once they completed 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 throughout the experiment to avoid possible data artifacts and false positives. 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 control and the experimental groups (Chittaro & Buttussi, 2015; Meng & Zhang, 2014; Tucker et al., 2018).

## Results

Welch's t-test and Student's t-test with a 5% significance level 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 also 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. Comparing the mean HR captured before, during, and after the gaming session makes it clear that 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,  $p=.99$ ). Similarly, there were no differences in the distribution of military training across the two groups (odds=0.93,  $p=.99$ ), on the use of underground parking (odds=1.31,  $p=.99$ ), frequency of gaming (odds=0.89,  $p<.646$ ), or familiarity with virtual reality (odds=0.68,  $p=.99$ ).

**Table 1**  
Continuous Demographic and Physiological Variables

Group	Control Group ( $n=10$ )	Experimental Group ( $n=7$ )	$t$ (df)	$p$	95% CI
-------	-----------------------------	---------------------------------	----------	-----	--------

Age	27.2	26.1	0.39(15)	.696	-4.608; 6.723
BMI <sup>a</sup>	23	24.8	-0.99(15)	.334	-5.794; 2.100
HR <sup>b</sup>	88.4	76.7	-63.892 (20888) ***	<.001	-12.11582; -11.39457

<sup>a</sup>Body Mass Index

<sup>b</sup>Heart Rate

\*\*\* Significant at  $p < .001$

### Game Behavioral Data

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). As demonstrated in Table 2, subjects in the control group made more pauses and turns, walked more and travelled faster, and took more time to find the exit. The lowest  $p$  value (.166) was found for the number of turns (Figure 4E), but this result was non-significant.

**Table 2**  
Game Performance Analytics Variables

Group	Control Group (n =10)	Experimental Group (n =7)	$t$ (df)	$p$	95% CI
Goal 1— Time taken to find the ambulance (sec)	24.0	28.0	-0.5(14.9)	.591	-22.34898; 13.20612
Goal 2— Time taken to find Exit door (sec)	206.0	168.0	0.55(15)	.587	-108.0815; 183.9958
Walked distance	571.0	483.0	0.62(15)	.541	-212.9697; 389.5697
Average speed	325.0	300.0	0.63(11)	.535	325.5000; 300.7143
Number of stops	53	25	1.14(11.6)	.274	-24.82213; 79.70785
Number of turns	297	163	1.4(15)	.166	-62.27264; 330.61550

### Heart Rate Sensor Data

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

**Table 3**  
Physiological Trends

Group	<i>Z</i>	<i>n</i>	<i>p</i>	<i>S</i>	<i>tau</i>
Experimental Group EG (n=7)	-13.094***	1183	<.001	-177559	-0.258
Control Group CG (n =10)	43.08***	2066	<.001	1348362	0.6402

\*\*\* Significant at  $p < .001$

### Psychometric Data

A psychometric game experience questionnaire (Ijsselsteijn et al., 2013; Poels et al., 2007) was issued to a larger sample of 24 users (for comparison HR data was only collected from 17 participants), 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. The questionnaire consisted of a core module for in-game (IG) experience and a postgame (PG) module. We used a nonparametric Kruskal-Wallis test (Table 4) for analyzing the questionnaire data. In the postgame module it revealed a statistically significant ( $p<.001$ ) difference in negative affect, with a mean rank of 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 ( $p=.001$ ) was 17.70 in the control group and



30.25 in the experimental group. It is important to note that the results are comparable between conditions in this study, but not to other similar studies.

In the core module measuring in-game experience, the results also indicated a statistically significant difference in positive affect ( $p=.007$ ) and challenge ( $p=.006$ ). Mean ranks for positive affect in the control group were 67.61 and 51.36 for the experimental group. Mean ranks for challenge were 50.07 for the control group and 66.92 for the experimental group respectively. These results indicate that in the experimental group, following the LED lights was possibly perceived as an extra challenge. For in-game experience “negative affect” and “tension and annoyance,” the  $p$  value was less than 0.001. These items are known to correlate in game experience questionnaires (Ijsselsteijn et al., 2013; Poels et al., 2007); therefore, similar results were 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.

**Table 4**  
Results for the In-Game and Post-Game Game Experience Questionnaire

Questionnaire Item	Group	$n^a$	Mean Rank <sup>b</sup>	$p$
In-Game Positive Affect	Control Group	55	67.61	.007**
	Experimental Group	62	51.36	
	Total	117		
In-Game Negative Affect	Control Group	44	36.23	<.001***
	Experimental Group	50	57.42	
	Total	94		
In-Game Tension/ Annoyance	Control Group	33	26.92	<.001***
	Experimental Group	39	44.60	
	Total	72		
In-Game Challenge	Control Group	55	50.07	.006**
	Experimental Group	62	66.92	
	Total	117		
Post-Game Negative Affect	Control Group	22	16.91	<.001***
	Experimental Group	26	30.92	

	Total	48		
Post-Game Tension	Control Group	22	17.70	.001***
	Experimental Group	26	30.25	
	Total	48		
<p><sup>a</sup> <i>n</i> refers to number of answers in each questionnaire item. Also, the questionnaire consists of a core module for in-game (IG) experience and a postgame (PG) module (Ijsselsteijn et al., 2013).</p> <p><sup>b</sup> Mean Ranks have been calculated using SPSS 25, they show how control and experimental groups rank against each other.</p> <p>*** Significant at <math>p &lt; .001</math></p> <p>** Significant at <math>p &lt; .01</math></p>				

## 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 the experimental condition by means of a signage system. Two groups of participants played the same game, 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 would outperform those in the control group (no lights). Moreover, we hypothesized that the HR data analysis would 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 users' in-game performances, self-reported psychometric questionnaires and physiological readings from a HR sensor.

### *Effect on group selection*

Standardized mean difference tests were used to compute the effect of demographic attributes between the control and experimental groups, and 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 & Isom, 2010; Kober & Neuper, 2011; Sandstrom, Kaufman & Huettel, 1998). 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 group of young adults (Head and Isom, 2010, p. 55). In addition, the same study reported significant gender deficits, with female participants being able to recall more landmarks than males. In another study with 48 undergraduate and graduate students, Sandstrom et. al (1998) observed that females participants perform better in navigation tasks when landmark information was present, but unlike male participants, 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. The “female participants showed a stronger increase in theta oscillations” when using landmarks as navigation aids (Kober and Neuper 2011, p.350). Additionally, motivated by Head and Isom (2010) in our search, 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, a t-test showed a significant variation in HR values, with participants in the control condition having larger means. Various factors can account for this variability, including outliers: some of these outliers might be owing to motion sickness felt by users, subjects with high BMI known to have high resting heart rates, and participants who engage in intense athletic activities who are known to have low resting heart rates.

### *Effect on behavioural outcomes*

Prior research has found strong correlations between wayfinding performance and visual sensory cues in VR installations (Feng et al., 2018; Radianti et al., 2020; Williams-Bell et al., 2015). In their experiment, Tang, Wu, and Lin (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 tested one hundred and seven subjects. The presence of either of the two exit sign types, according to their results, reduced the latency of players 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 coloured 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. Furthermore, their study also demonstrated that the placement of signs (wall vs floor) has a confounding effect on the performance of users. Players in the “floor condition” achieved better performance than those using the wall signage (Vilar et al., 2014).

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 (Cosma et al., 2016).

Our findings indicate that the performance and behavior of participants as captured by the in-game software is not statistically significant different, suggesting that the presence/absence of assistive LED lights had no significant direct effect in the completion of goals. 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 affordances (Gibson, 1977; 2015; Norman, 2013), 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 affordances to help users select from a number of potential navigation paths (Bonner & Epstein, 2017).

### *Effect on Heart Rate outcomes*

Results from the Mann-Kendall test show that HR values increased significantly for the control group, while they dropped for the experimental group. This finding 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 groups remained the same. They observed that the HR for both groups increased once the wayfinding simulation started, with the treatment group having overall a 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 that simulates 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).

Against this backdrop, our experiment held all gaming conditions constant besides the presence/absence of wayfinding signs. Moreover, analysis of self-reported psychometric data have shown that the experimental group experienced the game in drastically different ways (more negatively) than the control group, though we did not observe any important differences on how the two groups experienced simulation sickness, a factor which is known to affect HR variability (Cobb, Nichols, Ramsey, & Wilson, 1999; Holmes & Griffin, 2001). By extension, a possible explanation for the statistically significant difference in HR measurements seen in our results is a latent cognitive perception mechanism, an explanation for which prior literature exists. For instance, cognitive workload and 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), all while driving and shifting. An ECG signal showed that HR increased when participants were performing the dual activity (i.e., when driving). The authors

suggest that this can be explained by the competition between working memory components, prompted by the multiple cognitive tasks (Hidalgo-Muñoz et al., 2019).

Thus, we reason that in our experimental set up, the decreasing HR 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. On the other hand, the control group (no landmarks) exhibited a higher number of active electrodes in the left posterior inferior and superior regions of the brain; the authors speculate that the subjects in this group had a higher chance of disorientation, which “increased the re-wiring cost associated with such frequent cognitive demands” Sharma et al. (2017, p. 10).

## Limitations

This study suffers from limitations caused by small sample size. We however aimed at complementing this by collecting data from various sources. Participants used Oculus Rift S controllers and had free locomotion inside the simulation. Free locomotion in VR is known to cause cybersickness to some (); however all participants had the same controls. The recently criticized GEQ (Ijsselsteijn et al., 2013; Law, Brühlmann, & Mekler, 2018; Poels et al., 2007) was used to

assess the user experience of the participants. This questionnaire was chosen because it addresses general aspects of experience. Despite being originally hastily validated, GEQ has been widely used in numerous user studies. The recent de-validation of GEQ was conducted post-game long after the actual gameplay in an uncontrolled environment using Amazon mechanical Turk (IJsselsteijn et al., 2013; Law, et al., 2018). However, in this study we issued the questionnaire to the participants right after the experience. Still, the existing criticism needs to be mentioned as a limitation when interpreting the results of this study, specifically as we used the English version of the questionnaire with nonnative English speakers.

## **CONCLUSION**

The article presents the results of a VR experiment with twenty-four participants, simulating an underground parking wayfinding installation in Oculus Rift S HMD. Results indicate that participants in the control group traveled more, paused and turned more, and took more time to find the exit, but these differences were not found to be statistically significant. However, differences in HR outcomes between the two groups were found to be statistically significant, with subjects in the control condition displaying an increasing HR trend during simulation. This finding is aligned with prior results on the efficacy of landmarks and wayfinding affordances that have been shown to reduce cognitive demands and improve brain wiring connectivity (Hidalgo-Muñoz et al., 2019; Sharma et al., 2017).

Specifically, 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 that warrants further investigation is that there was more efficient cognitive processing in the experimental group, where wayfinding affordances were present, and that it was thus easier for the users to find the exit.



Simply put, assistive lights made pathfinding easier. This is in line with results from the psychometric self-reported data. Participants in the experimental condition also reported to have experienced a more negative gaming experience. This inconsistency can be attributed to subjects in this condition having fewer opportunities for exploring the scene, lack of a novelty effect, or not meeting their overly excited 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 assess user experience in VR. The inconsonance we have found between self-reported data and biometric measurements is noteworthy and warrants a re-inspection of many current findings in VR related research.

### **Acknowledgement**

This research has been conducted following the ethical requirements established by the Finnish Advisory Board on Research Integrity (TENK). The gathered material has been handled and informed consent from the participants was obtained in accordance with Finnish laws. Informed consent was acquired from each study participant.

### **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**

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Partly funded by European Commission grant PRINCE (grant 815362) and by Academy of Finland 6Genesis Flagship (grant 318927).

### **Bibliography**

- Alavesa, P., Pakanen, M., Niemelä, A., Huang, W., Väinämö, S., Haukipuro, L., Arhippainen, L., Ojala, T. (2018). Mobile Augmented Reality Client as a UX Method for Living Lab's User Involvement Tool. *22nd International Academic Mindtrek Conference, ACM*, New York, USA, 135-142. <https://doi.org/10.1145/3275116.3275147>
- Arias, S., La Mendola, S., Wahlqvist, J. et al., (2019). Virtual Reality Evacuation Experiments on Way-Finding Systems for the Future Circular Collider. *Fire Technology*, 55, 2319–2340. <https://doi.org/10.1007/s10694-019-00868-y>
- Bai, Y., Welk, G. J., Nam, Y. H., Lee, J. A., Lee, J. M., Kim, Y., et al., (2016). Comparison of consumer and research monitors under semistructured settings. *Medicine & Science in Sports & Exercise*, 48(1),151-158. <https://doi.org/10.1249/mss.0000000000000727>

- Bonner, M. F., Epstein, R. A. (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, S. M., An, H. S., Kang, S. K., Noble, J. M., Berg, K. E., Lee, J. M. (2017). Concurrent validity of wearable activity trackers under free-living conditions. *Journal of Strength and Conditioning Research*, *31*(4), 1097-1106. <https://doi.org/10.1519/jsc.0000000000001571>
- van Buuren, S., Wijnmalen, D. J. (2015). Measuring Psychosocial Impact of CBRN Incidents by the Rasch Model. *Journal of applied measurement*, *16*(3), 1-9.
- Chittaro, L., Buttussi, F. (2015). Assessing Knowledge Retention of an Immersive Serious Game vs. a Traditional Education Method in Aviation Safety, *IEEE Transactions on Visualization and Computer Graphics*, *21*(4), 529– 538. <https://doi.org/10.1109/tvcg.2015.2391853>
- Cho, D., Ham, J., Oh, J., Park, J., Kim, S., Lee, N.-K., & Lee, B. (2017). Detection of Stress Levels from Biosignals Measured in Virtual Reality Environments Using a Kernel-Based Extreme Learning Machine. *Sensors*, *17*(10), 2435. doi:10.3390/s17102435
- Christopher, K. (2014). *Port security management* (2nd ed.). CRC Press, Taylor Francis, 2014.
- Cobb, S. V., Nichols, S., Ramsey, A., & Wilson, J. R. (1999). Virtual reality-induced symptoms and effects (VRISE). *Presence: Teleoperators & Virtual Environments*, *8*(2), 169-186. <https://doi.org/10.1162/105474699566152>
- Cosma, G., Ronchi, E., Nilsson, D. (2016). Way-finding lighting systems for rail tunnel evacuation: A virtual reality experiment with Oculus Rift®. *Journal of Transportation Safety & Security*, *8*:101-117. <https://doi.org/10.1080/19439962.2015.1046621>
- Dondlinger, M. J., & Lunce, L. M. (2009). Wayfinding affordances are essential for effective use of virtual environments for instructional applications. *MERLOT Journal of Online Learning and Teaching*, *5*(3), 562-569. ISSN: 1558-9528.
- Feehan, L. M., Geldman, J., Sayre, E. C., Park, C., Ezzat, A. M., Yoo, J. Y., Hamilton, C. B., Li, L. C. (2018). Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data. *JMIR mHealth and uHealth*, *6*(8):e10527. <https://doi.org/10.2196/10527>
- Feng, Z., Vicente, A. González, R. A., Lovreglio, R. and Cabrera-Guerrero, G. (2018). Immersive Virtual Reality Serious Games for Evacuation Training and Research: A Systematic Literature Review. *Computers & Education*, *(127)*, 252–66. <https://doi.org/10.1016/j.compedu.2018.09.002>.
- Florea, C., Alavesa, P., Arhippainen, L., Pouke, M., Huang, W., Haukipuro, L., Väinämö, S., Niemelä, A., Cortés Orduña, M., Pakanen, M., & Ojala, T. (2019). *Extending a User Involvement Tool with Virtual and Augmented Reality*. IEEE Virtual Reality, IEEE VR, Osaka, Japan. <https://doi.org/10.1109/VR.2019.8798299>
- Föhr, T., Pietilä, J., Helander, E., Myllymäki, T., Lindholm, H., Rusko, H., & Kujala, U. M. (2016). Physical activity, body mass index and heart rate variability-based stress and recovery in 16 275 Finnish employees: a cross-sectional study. *BMC public health*, *16*, 701. <https://doi.org/10.1186/s12889-016-3391-4>

- Fokkema, T., Kooiman, T. J., Krijnen, W. P., van der Schans, C. P., de Groot, M. (2017). Reliability and validity of ten consumer activity trackers depend on walking speed. *Medicine & Science in Sports & Exercise*, 49(4), 793-800. <https://doi.org/10.1249/mss.0000000000001146>
- Giles, D., Draper, N., Neil, W. (2016). Validity of the Polar V800 heart rate monitor to measure RR intervals at rest. *European Journal of Applied Physiology*, 116(3), 563-571. <https://doi.org/10.1007/s00421-015-3303-9>
- Gibson, J.J. (1977) The Theory of Affordances. In R. E. Shaw & J. Bransford (Eds.), *Perceiving, Acting, and Knowing* (pp. 67-83). Lawrence Erlbaum Associates, Hillsdale, NJ. <https://doi.org/10.4324/9781315467931>
- Gibson, J.J. (2015). *The Ecological Approach to Visual Perception*. New York: Psychology Press, <https://doi.org/10.4324/9781315740218>
- Head, D., Isom, M. (2010). Age effects on wayfinding and route learning skills, *Behavioural Brain Research*, 209(1), 49-58. <https://doi.org/10.1016/j.bbr.2010.01.012>.
- Henriksen, A., Haugen, M. M, Woldaregay, A. Z., Muzny, M, Hartvigsen, G., Hopstock, L. A., Grimsgaard, S. (2018). Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables. *Journal of Medical Internet Research*, 20(3), e110. <https://doi.org/10.2196/jmir.9157>
- Hernández-Vicente, A., Santos-Lozano, A., De Cocker, K, Garatachea, N. (2016). Validation study of Polar V800 accelerometer. *Annals of Translational Medicine*, 4(15), 278. <https://doi.org/10.21037/atm.2016.07.16>.
- Hidalgo-Muñoz, A. R., Béquet, A. J., Astier-Juvenon, M., Pépin, G., Fort, A., Jallais, C., Tattegrain, H., Gabaude, C. (2019). Respiration and Heart Rate Modulation Due to Competing Cognitive Tasks While Driving, *Frontiers in Human Neuroscience*, (12), <https://doi.org/10.3389/fnhum.2018.00525>
- Hidayetoglu, M. L., Yildirim, K., & Akalin, A. (2012). The effects of color and light on indoor wayfinding and the evaluation of the perceived environment. *Journal of environmental psychology*, 32(1), 50-58. <https://doi.org/10.1016/j.jenvp.2011.09.001>
- Holmes, S. R., & Griffin, M. J. (2001). Correlation between heart rate and the severity of motion sickness caused by optokinetic stimulation. *Journal of Psychophysiology*, 15(1), 35. <https://doi.org/10.1027/0269-8803.15.1.35>
- IJsselsteijn, W. A., De Kort, Y. A. W., and Poels, K. (2013). The Game Experience Questionnaire. Eindhoven: *Technische Universiteit Eindhoven* 3–9.
- Jo, E., Lewis, K., Directo, D., Kim, M. J., Dolezal, B. A. (2016). Validation of biofeedback wearables for photoplethysmographic heart rate tracking. *Journal of Sports Science and Medicine*. 15(3),540–547.
- Johnson, D., Gardner, M. & Perry, R. (2018). Validation of two game experience scales: The Player Experience of Need Satisfaction (PENS) and Game Experience Questionnaire (GEQ). *International Journal of Human-Computer Studies*, 118, 38. doi: 10.1016/j.ijhcs.2018.05.003

- Kober, S. E. and Neuper, C. (2011). Sex differences in human EEG theta oscillations during spatial navigation in virtual reality, *International Journal of Psychophysiology*, 79(3), 347-355. <https://doi.org/10.1016/j.ijpsycho.2010.12.002>.
- Kostakos, P., Alavesa, P., Oppenlaender, J., Hosio, S. (2019). VR ethnography: a pilot study on the use of virtual reality “go-along” interviews in Google street view. *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia (MUM '19)*, ACM, New York, NY, USA, Article 53, 1–5. <https://doi.org/10.1145/3365610.3368422>
- Law, E. L.-C., Brühlmann, F., & Mekler, E. D. (2018). Systematic Review and Validation of the Game Experience Questionnaire (GEQ)—Implications for Citation and Reporting Practice. *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, 257–270. <https://doi.org/10.1145/3242671.3242683>
- Lynch, K. (1960). *The Image of the City*. Cambridge, MA: MIT/Harvard University press.
- May, M. (2004). Wayfinding, Ships and Augmented Reality. In: P. Andersen and L. Qvortrup (eds) *Virtual Applications* (pp. 2012-233). Springer, London. [https://doi.org/10.1007/978-1-4471-3746-7\\_10](https://doi.org/10.1007/978-1-4471-3746-7_10)
- Meng, F., Zhang, W. (2014). Way-finding during a fire emergency: an experimental study in a virtual environment. *Ergonomics*, 57(6), 816-827. <https://doi.org/10.1080/00140139.2014.904006>
- Mercer, K., Li, M., Giangregorio, L., Burns, C., Grindrod, K. (2016). Behavior change techniques present in wearable activity trackers: a critical analysis. *JMIR mHealth and uHealth*, 4(2), e40. <https://doi.org/10.2196/mhealth.4461>
- 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>
- Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic books. ISBN: 9780262525671
- Oculus. (2020). Retrieved from <https://www.oculus.com/rift-s/>
- 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. *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*, 688-692. <https://doi.org/10.1145/3267305.3267691>
- Poels, K., de Kort, Y. A. W., & IJsselsteijn, W. A. (2007). D3.3 : Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games. *Eindhoven: Technische Universiteit Eindhoven*.
- Polar. (2020) Polar A370 Fitness tracker with heart rate monitoring. Retrieved from <https://www.polar.com/en/products/sport/A370-fitness-tracker>
- Radianti, J., Majchrzak, T. A., Fromm, J., Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned,

- and research agenda. *Computers & Education*, 147. <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(3), 273–286. [https://doi.org/10.1162/pres\\_a\\_00049](https://doi.org/10.1162/pres_a_00049)
- Ronchi, E., Mayorga, D., Lovreglio, R., Wahlqvist, J., Nilsson, D. (2019). Mobile-powered head-mounted displays versus cave automatic virtual environment experiments for evacuation research. *Computer Animation and Virtual Worlds*, 30(6), e1873. <https://doi.org/10.1002/cav.1873>
- Ronchi, E., Nilsson, D., Kojić, S. et al., (2016). A Virtual Reality Experiment on Flashing Lights at Emergency Exit Portals for Road Tunnel Evacuation. *Fire Technology*, 52, 623–647. <https://doi.org/10.1007/s10694-015-0462-5>
- Ruddle, R. A. & Lessels, S. (2006). Three Levels of Metric for Evaluating Wayfinding. *Presence: Virtual and Augmented Reality*, 15(6), 637-654. <https://doi.org/10.1162/pres.15.6.637>
- Sandstrom, N. J., Kaufman, J., & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task, *Cognitive Brain Research*, 6(4), 351-360. [https://doi.org/10.1016/S0926-6410\(98\)00002-0](https://doi.org/10.1016/S0926-6410(98)00002-0).
- Sharma, G., Yash, K., Sushil, C., Vijander, S., Alok, M. P., & Varun, D. (2017). Influence of Landmarks on Wayfinding and Brain Connectivity in Immersive Virtual Reality Environment, *Frontiers in Psychology*, 8(1220) . <https://doi.org/10.3389/fpsyg.2017.01220>.
- Stahl, S. E., An, H. S., Dinkel, D. M., Noble, J. M., Lee, J. M. (2016). How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? *BMJ Open Sport & Exercise Medicine*, 2(1), 1–7. <https://doi.org/10.1136/bmjsem-2015-000106>.
- Su, F., Anthony, I. (2019). Reassessing CBRN Threats in a Changing Global Environment. *Solna, Sweden: Stockholm International Peace Research Institute*. Retrieved from [https://www.sipri.org/sites/default/files/2019-06/1906\\_cbrn\\_threats\\_su\\_anthony\\_0.pdf](https://www.sipri.org/sites/default/files/2019-06/1906_cbrn_threats_su_anthony_0.pdf)
- Tang, C.-H., Wu, W.-T., Lin, C.-Y. (2009). Using virtual reality to determine how emergency signs facilitate way-finding. *Applied ergonomics*, 40(4), 722–730. <https://doi.org/10.1016/j.apergo.2008.06.009>
- Terbizan, D. J., Dolezal, B. A., Albano, C. (2002). Validity of Seven Commercially Available Heart Rate Monitors. *Measurement in Physical Education and Exercise Science*, 6:4, 243-247. [https://doi.org/10.1207/S15327841MPEE0604\\_3](https://doi.org/10.1207/S15327841MPEE0604_3)
- Trend (2020) R package ‘trend’, January 13, 2020, Retrieved from <https://cran.r-project.org/web/packages/trend/trend.pdf>
- Tucker, A., Marsh, K. L., Gifford, T., Lu, X., Luh, P. B., Astur, R. S. (2018). The effects of information and hazard on evacuee behavior in virtual reality. *Fire Safety Journal*, 99, 1-11. <https://doi.org/10.1016/j.firesaf.2018.04.011>.
- Unity. (2020). Retrieved from <https://unity.com/>

- Unity API. (2019). Unity Scripting Reference Documentation - Version 2019.3. Retrieved from <https://docs.unity3d.com/ScriptReference/>
- Vilar, E., Rebelo, F., Noriega, P. (2014). Indoor Human Wayfinding Performance Using Vertical and Horizontal Signage in Virtual Reality. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 24(6), 601-615. <https://doi.org/10.1002/hfm.20503>
- W3C RQTF. (2020). Virtual Reality and Accessibility References by the Research Questions Accessibility Task Force (RQTF). Retrieved from [https://www.w3.org/WAI/APA/task-forces/research-questions/wiki/Virtual\\_Reality\\_and\\_Accessibility\\_References](https://www.w3.org/WAI/APA/task-forces/research-questions/wiki/Virtual_Reality_and_Accessibility_References)
- Wang, R., Blackburn, G., Desai, M., et al., (2017). Accuracy of Wrist-Worn Heart Rate Monitors. *JAMA Cardiology*, 2(1), 104–106. <https://doi.org/10.1001/jamacardio.2016.3340>
- Williams-Bell, F.M., Kapralos, B., Hogue, A. et al., (2015). Using Serious Games and Virtual Simulation for Training in the Fire Service: A Review. *Fire Technology*, 51(3), 553–584. <https://doi.org/10.1007/s10694-014-0398-1>
- Zhao, C. M., Lo, S. M., Zhang S. P., Liu, M. (2009). A Post-Fire Survey on the Pre-Evacuation Human Behavior. *Fire Technology*, 45(1). <https://doi.org/10.1007/s10694-007-0040-6>