Comparing Visual Search between Physical Environments and VR

Floris van den Oever[*]
Department of Psychosocial Science
University of Bergen

Valentina Gorobets[†]
Institute of Machine Tools and Manufacturing
ETH Zurich

Bjørn Sætrevik[‡]
Department of Psychosocial Science
University of Bergen

Morten Fjeld[§]
Department of Information Science and Media Studies
University of Bergen

Andreas Kunz[¶]
Institute of Machine Tools and Manufacturing
ETH Zurich

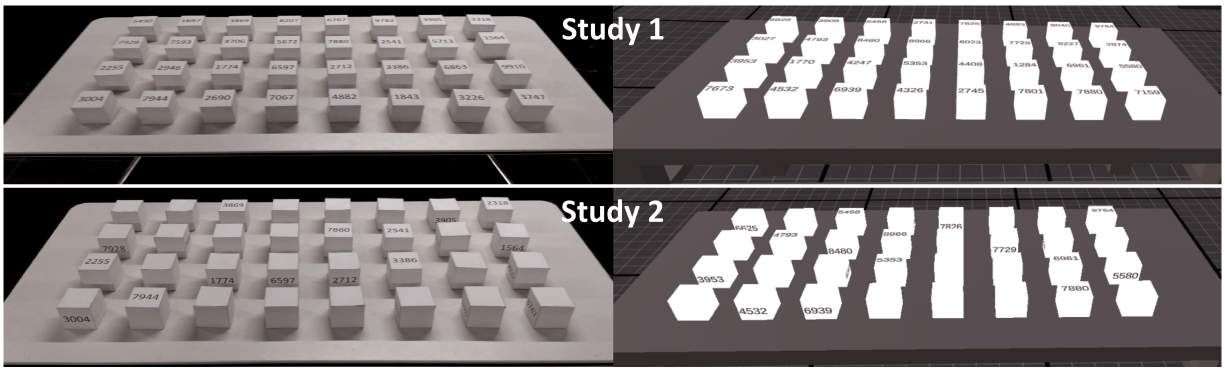

Figure 1: Study setups. From left to right and top to bottom: Study one in the physical environment, Study one in VR, Study two in the physical environment, and Study two in VR.

Abstract

Virtual reality (VR) is increasingly used in research and practice under the assumption that human behavior in VR generalizes to behavior in physical environments. This assumption lacks extensive testing, thus risking unexpected human behavior when working with VR. Therefore, we believe it would be beneficial to evaluate whether humans perform tasks in VR in similar ways as in physical environments. In the current study, we compared the performance on visual search tasks in physical environments and in VR. The participants (n = 29) performed search tasks while standing still and search tasks that required moving around in both physical environments and in VR. We measured search speed, accuracy, workload, and cognitive absorption and analyzed them with Bayesian t-tests and Bayesian factorial ANOVAs. Our results provide weak to moderate evidence that all were equal in VR as in the physical environments, even when controlling for VR experience and personal innovativeness. Our findings provide some evidence for the assumption that VR can simulate or replace visual search tasks in physical environments. Because our study focused on the narrow topic of visual search in a very controlled environment, we advise further scrutiny of the underlying assumptions of VR use, for example, by studying more naturalistic search scenarios, the role of memory, and interaction with objects.

Index Terms: H.5.2 [User Interfaces]: User Interfaces—Graphical user interfaces (GUI); H.5.m [Information Interfaces and Presentation]: Miscellaneous

[*]e-mail: floris.oever@uib.no
[†]e-mail: gorobets@iwf.mavt.ethz.ch
[‡]e-mail: bjorn.satrevik@uib.no
[§]e-mail: morten.fjeld@uib.no
[¶]e-mail: kunz@ethz.ch

1 Introduction

Virtual Reality (VR) is used in various applications, such as remote operation of maritime vessels [8, 32]; remote operation of aerial vehicles [15, 16, 30]; air traffic control [4, 10]; maritime traffic control [53]; remote support during maintenance and inspection [33]; training [25, 36, 40, 44]; architecture and design [2, 28, 44]; psychological therapy, [35]; and surgery, [3, 12]. Since the use of VR is increasing, its ecological validity and applicability for these operations should be evaluated to ensure its feasibility, safety, and effectiveness. For example, Harris et al. point out the inadequate understanding of how VR impacts cognitive factors [23].

In performing industrial operations such as these, visual search plays an important role [14]. Therefore, it is important to investigate if visual search performance in virtual environments correlates with physical environments. In the present study, we adopt Burack et al.'s [7] definition of visual search: *'Visual search is a goal-oriented activity that occurs regularly in daily life and involves the active scanning of the environment in order to locate a particular target among irrelevant non-targets, or distractors'*.

1.1 Related Work

The previous section shows the prevalence of direct use and evaluation of VR for operations involving visual search. We, however, are interested in more fundamental empirical knowledge on visual search in VR to ground VR research and innovation. In this section we discuss related work on visual search in VR, the comparison of visual search between VR and 2D displays, and the comparison of visual search related tasks between VR and physical environments.

Visual search tasks have been studied in VR to understand visual search in physical environments; for example, visual search in traffic while fatigued [52]; the effects of distractions during visual search [38]; the effects of impairments on visual search [6, 9]; or general visual search constructs, such as the relative importance of central and peripheral vision [11] and the search initiation effect [5]. Like the implementation of VR in practice, these studies lean on the assumption that visual search behavior and performance in VR can be mapped to physical environments. As such, grounding this assumption in fundamental empirical findings is valuable to support further use of VR for visual search in research as well as in practice.

One approach towards increasing the validity of theories about visual search is that classic visual search experiments on 2D displays are being reproduced in VR. While knowledge generated in 2D display experiments has been generalized to phenomena in physical environments, the caveat of uncertain and probably low ecological validity has been mentioned [22, 24]. To evaluate the concurrent validity of 2D computer experiments and 3D VR experiments, visual search tasks on 2D displays and in VR have been compared. For example, Hadnett-Hunter et al. [22] replicated experiments on feature search, wide field-of-view visual search, and eccentricity effects. Similarly, Li et al. [34] investigated the role of spatial memory in visual search in both 2D and 3D environments. While results from Hadnett-Hunter et al. [22] are currently not available, Li et al. [34] found that search behavior on 2D display and in VR environments was largely comparable, although there was some behavioral difference signaling greater use of memory in VR. In like manner, Figueroa et al. [18] compared search efficiency on 2D displays and in VR, finding that visual search was faster and more accurate in VR. All three studies assumed higher ecological validity of VR than 2D displays and advocated for using VR for similar studies.

Differences and similarities in behavior and performance between physical environments and VR in tasks related to visual searching have been researched too. For example, Richardson et al. [42] compared the learning of a building layout using a map, direct experience, and VR, with results indicating that participants learned the overall layout worst in VR but performed equally well on learning the layout of landmarks. On the other hand, Kuliga et al. [31] investigated the validity of VR as a representation of physical environments, finding that study participants experienced a VR and a real building similarly, indicating that VR has a strong potential to be used as an empirical research tool. Congruently, Feldstein et al. [17] found similar distance estimation in physical environments and VR, which supports VR's applicability for visual tasks, given that important factors such as eye height are equal in physical environments and VR. In contrast, Kelly et al. [29] and Peillard et al. [39] found differences in distance perception between physical environments and VR, and suggested methods to reduce this difference. VR as a proxy for physical environments has been studied in the area of social attention, suggesting VR may elicit modes of information processing similar to those in physical environments but calling for direct comparisons of behavior in physical environments and VR [43]. Based on a literature review, Wilson & Soranzo [50] stated that higher levels of visual fidelity and immersion in VR do not necessarily elicit realistic psychological responses, suggesting that physical and psychological side effects from VR exposure should be taken into account when generalizing from VR to physical environments.

1.2 Research Gap and Study Aim

The apparent difference in visual search performance and behavior between 2D displays and VR, in combination with the previous generalization of 2D display based experiment results to physical environments, motivated us to check the assumption of the generalizability of VR experiment results to physical environments. Although the face validity of VR being more generalizable to physical environments than 2D displays is high, the assumption of generalizability remains understudied. Furthermore, the higher similarity of VR to physical environments opens the door to more exact comparison of visual search behavior and performance between VR and physical environments than is possible with 2D displays.

In our search for relevant literature, we found no research specifically focused on a direct comparison of visual search performance in 3D between physical environments and VR, even though researchers call for the evaluation of this novel methodological tool [22, 43, 50]. Concurrently, assumptions about ecological validity of experiments in general are under critique in fields like Psychology [26], while at the same time VR is being implemented to replace [8, 10, 15] and simulate [2, 3, 36] visual search tasks. Based on these considerations, we believe it is important to map the ecological validity of VR, thus paving the way for its feasible, safe, and effective implementation. Therefore, we empirically compared visual search performance between physical environments and VR. As such, our study contributes by testing the ecological validity of VR in visual search tasks, and demonstrating a method for comparing visual tasks between physical environments and VR, which can be applied to other activities than visual search.

1.3 Hypotheses

To compare visual search performance between physical environments and VR we examined four hypotheses and carried out additional exploratory analyses.

Hypothesis 1: People are equally fast in physical environments and VR. Researchers and practitioners seem to assume people behave and perform similarly in physical environments and VR, so we base our hypothesis on that assumption, even though Figueroa et al. [18] and Li et al. [34] found that search speed was higher in VR than in 2D, suggesting it may be even higher in physical environments.

Hypothesis 2: People are equally accurate in physical environments and VR. Both VR users and practitioners seem to assume people behave and perform similarly in physical environments and VR, so we base our hypothesis on that assumption, even though Figueroa et al. [18] and Li et al. [34] found that accuracy was higher in VR than in 2D, suggesting it may be even higher in a physical environment.

Hypothesis 3: People experience more workload in VR than in physical environments. We expect this because of the added complexity VR brings to the task over the task in a physical environment. When using VR, people have to adjust and operate a VR headset and experience a virtual environment while also being in a physical environment they do not see. On top of that, the novelty of VR may cause people to feel a higher workload [41].

Hypothesis 4: People experience more cognitive absorption in a physical environment than in VR. Cognitive absorption is linked to agency [20], which we expect to be higher in a physical environment. On top of that, the novelty of VR may distract people which may result in less cognitive absorption [41].

Exploratory analysis: Differences in speed and accuracy may be moderated by the VR experience and innovativeness, meaning that people with little or no VR experience and low innovativeness may have a larger difference in performance between physical environments and VR than those with more prior experience and higher innovativeness [41]. Therefore, we explored possible moderating effects of these variables.

2 Method

A within-subjects, randomized, controlled trial experiment following the CONSORT statement [45], as depicted in Fig. 2, was conducted. Ethical approval by an ethics committee was not requested because we saw no risks in participating, data were anonymized, and participants gave informed consent. Data collection took place in week 14 of 2022.

2.1 Participants

Our study was conducted with 29 participants. The participants were mainly recruited from the university faculty, from the student body, and some from outside the university. Participants were included if they had normal or corrected-to-normal vision and color vision, and could walk. Demographic characteristics of the participants are portrayed in Table 1.

We conducted a sample size calculation based on the planned analysis for the comparison of search speed. We calculated the sample size with the goal to have a power of 0.9 and type 1 error rate allowance of .05. We based our assumptions of group means and standard deviations on the results of two previous studies conducting similar analyses: based on the data of Figueroa et al. [18] who compared search tasks between VR and 2D display, our minimum sample size was 17; based on the data of Frederiksen et al. [19], who compared secondary task reaction time between immersive VR and conventional VR, our minimum sample size was 12. Because our participants carried out relatively few tasks for a visual search experiment and we wanted to have room for missing or excluding data, we aimed for a minimal sample of 20 participants, with the stopping rules of 30 participants and a fixed end date.

Table 1: Sample characteristics.

Characteristic	Statistic
Gender	
Man	23
Woman	6
Age in years, median (range)	24 (19, 49)
Occupation	
Student	22
PhD student or candidate	4
Other	3
Vision	
Normal	18
Corrected to normal	11
VR experience	
None	5
0 - 5 hours	15
5 - 20 hours	2
20 - 100 hours	3
>100 hours	4
Personal innovativeness, mean (standard deviation)	5.2 (1.3)

2.2 Experiment Design

Our experiment consisted of two studies. In Study one participants carried out search tasks while standing still; in Study two participants carried out search tasks for which they had to move around. In both studies, participants saw similar setups in VR and a physical environment (see Fig. 1). Each setup consisted of a 70cm high table with a surface of 120×60cm. Thirty two white cubes with edges of 7.8cm were located on the table. They were spread out in four rows and eight columns, with 8.2cm space between the cubes. Each cube had a number between 1000 and 9999 in large black font randomly assigned to it. There were no numbers that risked misinterpretation when reading from another angle than up front, such as 9969. In Study one, all numbers on the cubes faced upwards; In Study two, numbers were randomly located on different visible sides of the cubes (see Fig. 1) so participants had to walk around the table to find search targets. To prevent learning effects, randomization of numbers was done for all four setups (Study one in the physical environment, Study one in VR, Study two in the physical environment, and Study two in VR). In like manner randomization of orientation of the cubes in Study two was done for both setups (the physical environment and VR) to prevent learning effects. We choose white cubes with four digit numbers in a rectangular array to prevent learning effects based on scene grammar [13] and memorization.

In both studies, participants were asked to consecutively find 12 cubes with given numbers in both the physical environment and VR. Participants received audio instructions through headphones. Instructions started with the first click of a laser pointer, giving a randomly selected number to find among the cubes on the table. Once a participant found the target cube, they pointed at it with the laser and clicked a button to indicate they found it. On clicking, they heard the next number to find. To familiarize participants with the audio format of the instructions, they listened to two examples of the search tasks at the start of the experiment.

Fig. 2 shows the flow of the experiment. At the start of the experiment, each participant filled in the initial questionnaires. Then, participants were randomly assigned to start with the physical environment condition or VR condition for the first study where all cubes were facing upwards. After each condition, they filled in questionnaires. Once a participant was assigned to start the first study in VR or the physical environment, they also started the second study in the corresponding environment.

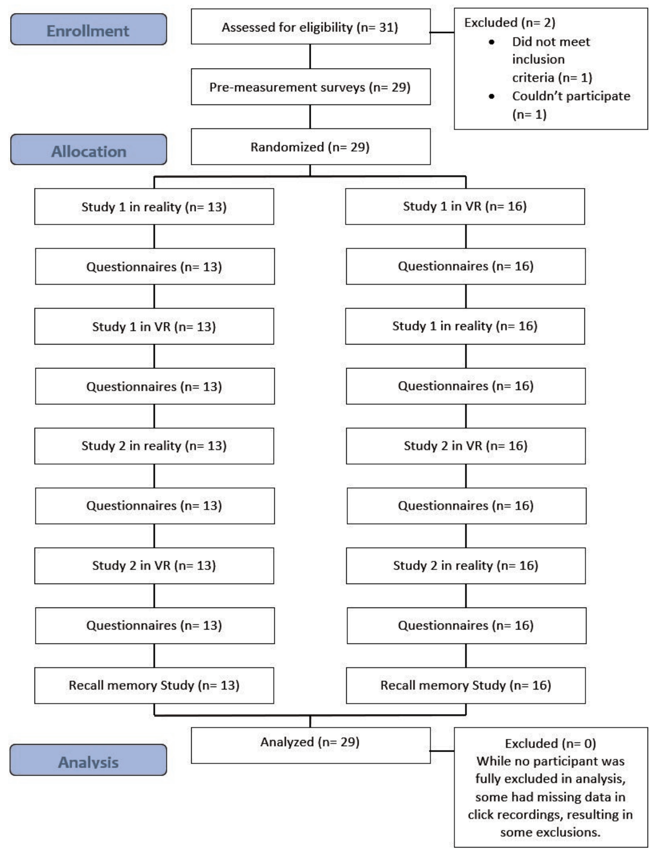

Figure 2: Procedure following the CONSORT statement [45].

2.3 Technical Setup

In the physical environment setup, each cube was printed on a sheet of regular office paper and folded. The virtual environment was developed in Unity and implemented with an HTC Vive Pro headset and one controller. In both the physical and the virtual environment, we captured timestamps when the button of the laser pointer was pressed. The laser pointer used in the physical environments was a Logitech R400. In the virtual setup (Study one), the HTC Vive Controller was used as the laser pointer, casting a laser ray in VR. In both setups, the next audio instruction was initiated by pressing the laser casting button. The audio instructions were created with a text-to-speech converter [1]. For all randomizations, a random number generator was used [21].

[1] https://www.ttsmp3.com

2.4 Outcome

Descriptive data were collected with questionnaires. The primary outcomes were time measurements, operationalized as the time participants needed to perform each search task in milliseconds (ms) since the last click. This includes the time needed to listen to the search task. To measure time since the last click we tracked clicks, self-coded in VR and with Mini Mouse Macro Recorder in the physical environment [47].

Secondary outcomes were accuracy, workload, and cognitive absorption. We measured accuracy after Study two, operationalized as the proportion of trials in which participants correctly identified the search target. To measure accuracy, we visually observed where the participants pointed when they clicked. We only measured accuracy for the second study because we expected too few errors when all the numbers were constantly in view. We measured workload with a 10-point Likert-scale, unweighted SIM-TLX [23]. We measured cognitive absorption with a 7-point Likert-scale, unweighted cognitive absorption questionnaire [1]. Lastly, to check the moderating effects of experience and innovativeness, we measured innovativeness with the Personal Innovativeness in IT questionnaire [1, 46] and asked how many hours of experience participants had with VR with the answer options: this is my first time using AR, 0-5 hours, 5-20 hours, 20-100 hours, and >100 hours.

2.5 Analysis

To test hypotheses 1-4, we used Bayesian t-tests. To compare search speed, we took the average time of the 12 search tasks. We pairwise excluded search speed data of participants with lost tracking data in one of the conditions. Outlying individual search task times with known causes were excluded before taking the averages of search task time (e.g., accidental double clicks, headset problems, or interruptions). We chose to not exclude outliers not caused by such technical issues, because we believe that they were part of the phenomena we were interested in [37]. We used Bayesian paired samples t-tests to compare search speed per search task. To compare accuracy, workload, and cognitive absorption between physical environments and VR we took average scores of each variable and conducted Bayesian paired samples t-tests.

For the exploratory analyses, we used Bayesian factorial ANOVAs. We analyzed the moderating effects of the factors VR experience and personal innovativeness on speed and accuracy. To make subgroups larger, we made both variables binary. For VR experience, we grouped participants with no or 0-5 hours of VR experience in "inexperienced" and participants with 5 or more hours of experience in "experienced". For personal innovativeness, we grouped participants scoring higher than the median in "more innovative" and participants scoring lower than the median in "less innovative".

We used Bayesian versions of classic parametric tests (t-tests and ANOVAs). We used Bayesian statistics as opposed to frequentist statistics so we could obtain evidence of the absence of effects [49]. We used default priors because related literature either focused on slightly different topics or did not provide sufficient insights in their statistics for us to build priors on. We used parametric tests (t-tests and ANOVAs) to analyze raw reaction time data, which is a frequently applied method [37, 51]. All statistics were carried out using JASP [27].

3 Results

In this section we present the results per hypothesis with regard to speed, accuracy, workload, and cognitive absorption. We also present exploratory analyses of the moderating effects of personal innovativeness and VR experience on search speed and accuracy.

3.1 Difference in Speed between Physical Environments and VR

Our findings support the first hypothesis. There is moderate evidence that people are equally fast in physical environments and VR based on Study one (Bayes Factor_{01} = 4.485) as well as based on Study two (BF_{01} = 3.692). This means that the observed data are approximately 4.5 and 3.7 times more likely to occur if people are equally fast in physical environments and VR than if there is a speed difference. The error percentage of the numerical algorithms used to obtain the results in Study one was 0.027% and in Study two was 0.029%, which indicates great stability of the algorithms. Fig. 3 shows arrow plots of the difference in speed in Study one and two with 95% credible intervals. In order to assess the robustness of the Bayes Factor (BF) to our prior specifications, Fig. 4 shows BF_{01} as a function of the prior width r. Across a wide range of widths the BF does not change our conclusion and can thus be seen as robust.

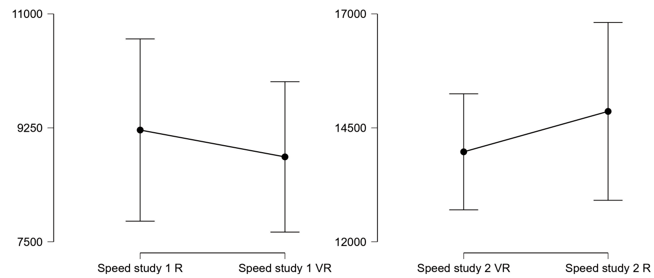

Figure 3: Relative search speed in ms in Study one (left) and Study two (right).

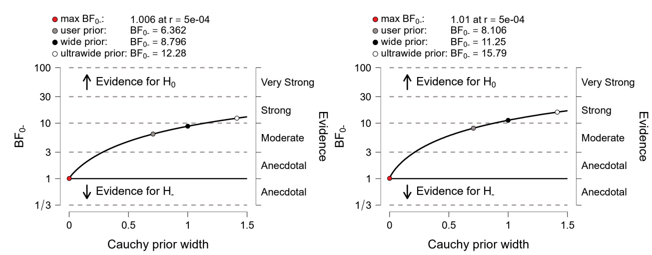

Figure 4: BF robustness check of speed in Study one (left) and in Study two (right).

3.2 Difference in Accuracy between Physical Environments and VR

Our findings support hypothesis two. There is weak evidence that people are equally accurate in physical environments and VR (BF_{01} = 1.571). This means that the observed data are approximately 1.5 times more likely to occur if people are equally accurate in physical environments and VR than if there is a difference in accuracy. The error percentage of the numerical algorithms used to obtain the results was 0.027% , which indicates great stability of the algorithms. However, the Bayes factor (BF) is not very robust. Fig. 5 shows on the left an arrow plot of the difference in accuracy between physical environments and VR. The BF robustness check on the right shows the BF would have been different with different prior distributions. There might be a ceiling effect; the actual difference in accuracy may be dampened by most participants correctly identifying all or most search targets [48]. As such, our finding does not warrant an all-or-none acceptance of hypothesis two.

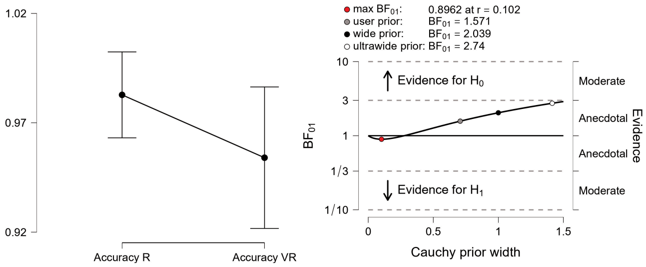

Figure 5: Relative accuracy in finding search targets (left) and BF robustness check (right).

3.3 Difference in Workload between Physical Environments and VR

Our findings do not support hypothesis three. There is weak evidence that people experience equal workload in physical environments and VR based on Study one (BF_{01} = 2.522), and moderate evidence based on Study two (BF_{01} = 4.584). This means that the observed data are approximately 2.5 and 4.6 times more likely to occur if people experience equal workload in physical environments and VR than if they experience more workload in VR. The error percentage of the numerical algorithms used to obtain the results in Study one was <0.001% and in Study two was 0.034%, which indicates great stability of the algorithms. Fig. 6 shows arrow plots of the difference in workload in Study one and two with 95% credible intervals. Fig. 7 shows that across a wide range of widths the BF does not change our conclusion and can thus be seen as robust. Generally, participants experienced a moderately low workload in both, Study one (M = 2.879, SD = 1.169) and two (M = 3.138, SD = 1.429) on a 10-point Likert scale.

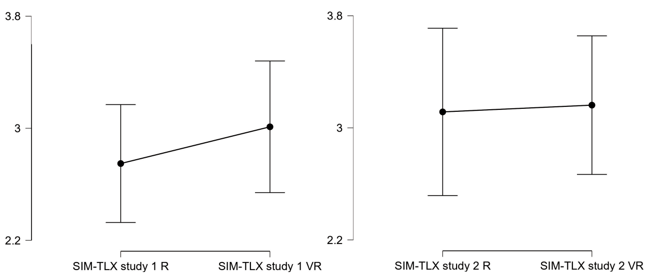

Figure 6: Relative workload in Study one (left) and Study two (right).

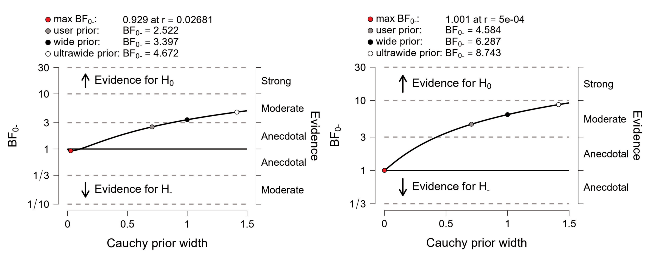

Figure 7: BF robustness check of workload in Study one (left) and in Study two (right).

3.4 Difference in Cognitive Absorption between Physical Environments and VR

Our findings do not support hypothesis four. There is strong evidence that people feel equally cognitively absorbed in physical environments and VR based on Study one (BF_{01} = 13.673) and Study two (BF_{01} = 10.446). This means that the observed data are approximately 13.6 and 10.4 times more likely to occur if people feel equally cognitively absorbed in physical environments and VR than if they feel more absorbed in physical environments. The error percentage of the numerical algorithms used to obtain the results in Study one was 0.086% and in Study two was <0.001%, which indicates great stability of the algorithms. Fig. 7 shows that across a wide range of widths the BF does not change our conclusion and can thus be seen as robust. Fig. 9 shows arrow plots of the difference in workload in studies one and two with 95% credible intervals. Generally, participants experienced high cognitive absorption in studies one (M = 5.931, SD = 0.999) and two (M = 5.834, SD = 1.015) on a 7-point Likert scale.

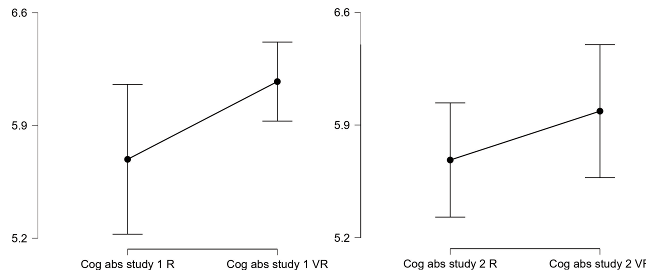

Figure 8: Relative cognitive absorption in Study one (left) and Study two (right).

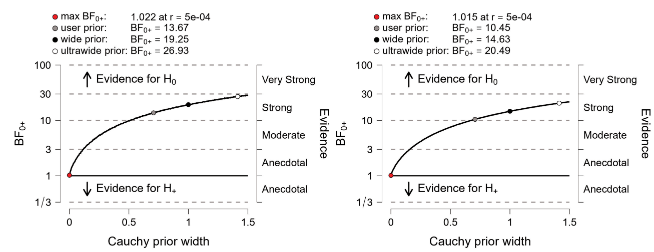

Figure 9: BF robustness check of cognitive absorption in Study one (left) and in Study two (right).

3.5 The Moderating Effects of Innovativeness and VR Experience

In addition to hypothesis testing, we conducted exploratory analyses of the moderating effects of the VR experience and personal innovativeness. For this we used Bayesian factorial ANOVAs for search speed and accuracy with the fixed factors of condition (physical environment or VR), VR experience, and personal innovativeness.

For speed, the Bayesian factorial ANOVAs indicated that the data were 2.667 times more likely in Study one and 1.37 times more likely in Study two to occur under the null model than under the second most predictive model. This indicates weak evidence that there is no moderating effect of VR experience nor of personal innovativeness. The error percentage of the numerical algorithms used to obtain the results was 0.009% in both Study one and two, which indicates great stability of the algorithms. For accuracy, the Bayesian factorial ANOVA indicated that the data were 1.379 times more likely to occur under the null model than under the second most predictive model. The error percentage of the numerical algorithms used to obtain the results was 0.009%, which indicates great stability of the algorithms. This indicates weak evidence that there is no moderating effect of VR experience nor of personal innovativeness.

4 Discussion

We compared the performance of people on a visual search task in physical environments and VR. The participants (n = 29) performed two types of search tasks in VR and in physical environments: in Study one, they stood still searching objects on a table, in Study two they had to walk around the table to find search targets. We measured search speed, accuracy, workload, and cognitive absorption, and analyzed these using Bayesian statistics. Our results provide weak to moderate evidence that search speed, workload, accuracy, and cognitive absorption are equal in VR as in physical environments, even when controlling for VR experience and personal innovativeness.

Our findings are in line with previous assumptions [11, 38] and findings [17, 31] that VR can represent physical environments, thus providing some support for the assumption that VR can be used to replace [8, 10, 15] and simulate [2, 3, 36] visual search tasks. As such we give some insight towards an answer to the call for the evaluation of this novel tool [22, 43, 50].

Several limitations should be taken into consideration when interpreting our findings. First, we used uninformed priors for our Bayesian tests. We considered using informed priors but could not find satisfying previous research results to inform them. Second, as can be seen in Fig. 8, there was an initially surprising find that our participants experienced higher cognitive absorption in VR than in physical environments. If we had hypothesized that people were more cognitively absorbed in VR, we could have collected evidence for a difference. This might be related to the novelty effect of VR [41]. Third, due to our sample size, our exploratory analyses on moderating effects have limited power. Perhaps stronger effects could be observed in a more extensive study. Fourth, we measured memory performance but found a large confounding variable due to our experiment setup, therefore we decided not to analyze memory performance. Fifth, we used t-tests and ANOVAs to analyze reaction time data. These are two widely used possible methods [37], but there is a chance that other methods would have given stronger evidence. Lastly, we focused on performance measures, but including behavioral measures such as eye-tracking may give extra insight into similarities and differences in human behavior between physical environments and VR.

Our findings support the assumption that VR can simulate or replace visual search tasks in research and practice, thus endorsing its implementation for such purposes. Because our study focused on the narrow topic of visual search in a very controlled environment, we advise further scrutiny of the underlying assumptions of VR use. For such research, we suggest using behavioral measures such as eye- and movement-tracking; studying more naturalistic search scenes; studying the role of memory; evaluating other existing search tasks; including interaction with objects; and isolating VR characteristics that are suspected to cause differences in behavior and performance such as limited field-of-view [22] and being able to move through virtual objects. Our analysis files can be used to inform hypotheses and prior distributions, available on OSF.io. Studies like ours can demonstrate whether VR can safely and effectively replace or simulate visual search tasks in both research and practice, so that safe and effective implementation can be achieved.

Acknowledgments

The authors wish to thank Marco Ketzel for helping with the VR development and Joy Gisler for helping with the ethical approval. This work was supported in part by the mobility fund of the University of Bergen.

References

[1] R. Agarwal and E. Karahanna. Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly*, 24(4):665–694, 2000.

[2] K. Aylward, J. Dahlman, K. Nordby, and M. Lundh. Using operational scenarios in a virtual reality enhanced design process. *Education Sciences*, 11(8), 2021. doi: 10.3390/educsci11080448

[3] A. Ayoub and Y. Pulijala. The application of virtual reality and augmented reality in oral & maxillofacial surgery. *BMC Oral Health*, 19(1), Nov. 2019. doi: 10.1186/s12903-019-0937-8

[4] J. Begli. Supporting small airports using virtual reality, Apr 2020.

[5] J. Beitner, J. Helbing, D. Draschkow, and M. L.-H. Võ. Get Your Guidance Going: Investigating the Activation of Spatial Priors for Efficient Search in Virtual Reality. *Brain Sciences*, 11(1):44, Jan. 2021. doi: 10.3390/brainsci11010044

[6] C. R. Bennett, E. S. Bailin, T. K. Gottlieb, C. M. Bauer, P. J. Bex, and L. B. Merabet. Virtual Reality Based Assessment of Static Object Visual Search in Ocular Compared to Cerebral Visual Impairment. In M. Antona and C. Stephanidis, eds., *Universal Access in Human-Computer Interaction. Virtual, Augmented, and Intelligent Environments*, pp. 28–38. Springer International Publishing, Cham, 2018.

[7] J. Burack, T. Dawkins, J. Stewart, H. Flores, G. Iarocci, and N. Russo. "the mysterious myth of attention deficit..." revisited: A discussion of how the developmental approach is transforming the understanding of intellectual disability. *International Review of Research in Developmental Disabilities*, 42(C):147–177, Jan. 2012. doi: 10.1016/B978-0-12-394284-5.00006-1

[8] H.-C. Burmeister, R. Grundmann, and B. Schulte. Situational awareness in ar/vr during remote maneuvering with mass: The tug case. In *Global Oceans 2020: Singapore – U.S. Gulf Coast*, pp. 1–6, 2020. doi: 10.1109/IEEECONF38699.2020.9389455

[9] H. Chow-Wing-Bom, T. Dekker, and P. Jones. Using Virtual Reality [VR] to assess the effects of asymmetric vision loss on visual search performance. *Journal of Vision*, 18(10):657, Sept. 2018. doi: 10.1167/18.10.657

[10] M. Cordeil, T. Dwyer, and C. Hurter. Immersive solutions for future air traffic control and management. In *Proceedings of the 2016 ACM Companion on Interactive Surfaces and Spaces*, ISS '16 Companion, p. 25–31. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/3009939.3009944

[11] E. J. David, J. Beitner, and M. L.-H. Võ. The importance of peripheral vision when searching 3d real-world scenes: A gaze-contingent study in virtual reality. *Journal of Vision*, 21(7), Jul 2021. doi: 10.1167/jov.21.7.3

[12] M. R. Desselle, R. A. Brown, A. R. James, M. J. Midwinter, S. K. Powell, and M. A. Woodruff. Augmented and virtual reality in surgery. *Computing in Science & Engineering*, 22(3):18–26, May 2020. doi: 10.1109/mcse.2020.2972822

[13] D. Draschkow and M. L.-H. Võ. Scene grammar shapes the way we interact with objects, strengthens memories, and speeds search. *Scientific Reports*, 7(1), Nov 2017. doi: 10.1038/s41598-017-16739-x

[14] M. P. Eckstein. Visual search: A retrospective. *Journal of Vision*, 11(5):14–14, 12 2011. doi: 10.1167/11.5.14

[15] J. M. Ernst and L. Ebrecht. Virtual cockpit instruments on head-worn displays for helicopter offshore operations in confined areas. In *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC)*, pp. 1–10, 2020. doi: 10.1109/DASC50938.2020.9256681

[16] J. M. Ernst, L. Ebrecht, and S. Schmerwitz. Virtual cockpit instruments displayed on head-worn displays – capabilities for future cockpit design. In *2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC)*, pp. 1–10, 2019. doi: 10.1109/DASC43569.2019.9081733

[17] I. T. Feldstein, F. M. Kölsch, and R. Konrad. Egocentric distance perception: A comparative study investigating differences between real and virtual environments. *Perception*, 49(9):940–967, Sep 2020. doi: 10.1177/0301006620951997

[18] J. C. M. Figueroa, R. A. B. Arellano, and J. M. E. Calinisan. A Comparative Study of Virtual Reality and 2D Display Methods in Visual Search in Real Scenes. In D. N. Cassenti, ed., *Advances in Human Factors in Simulation and Modeling*, pp. 366–377. Springer International Publishing, Cham, 2018.

[19] J. G. Frederiksen, S. M. Sørensen, L. Konge, M. B. Svendsen, M. Nobel-Jørgensen, F. Bjerrum, and S. A. Andersen. Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: A random-

ized trial. *Surgical Endoscopy*, 34(3):1244–1252, Jun 2019. doi: 10.1007/s00464-019-06887-8

[20] H. Freude, C. Reßing, M. Müller, B. Niehaves, and M. Knop. Agency and body ownership in immersive virtual reality environments: A laboratory study. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020.

[21] M. Haahr. True random number service.

[22] J. Hadnett-Hunter, E. O'Neill, and M. J. Proulx. Contributed session ii: Visual search in virtual reality (vsvr): A visual search toolbox for virtual reality. *Journal of Vision*, 22(3):19, Feb 2022. doi: 10.1167/jov.22.3.19

[23] D. Harris, M. Wilson, and S. Vine. Development and validation of a simulation workload measure: the simulation task load index (SIM-TLX). *Virtual Reality*, 24(4):557–566, Dec. 2019. doi: 10.1007/s10055-019-00422-9

[24] J. Helbing, D. Draschkow, and M. L.-H. Võ. Search superiority: Goal-directed attentional allocation creates more reliable incidental identity and location memory than explicit encoding in naturalistic virtual environments. *Cognition*, 196:104147, 2020. doi: 10.1016/j.cognition.2019.104147

[25] C. Hirt, M. Spahni, Y. Kompis, D. Jetter, and A. Kunz. Virtual reality training platform for a computer numerically controlled grinding machine tool. *International Journal of Mechatronics and Manufacturing Systems*, 14(1):1–17, 2020. doi: 10.1504/IJMMS.2021.115460

[26] G. A. Holleman, I. T. C. Hooge, C. Kemner, and R. S. Hessels. The 'real-world approach' and its problems: A critique of the term ecological validity. *Frontiers in Psychology*, 11, 2020. doi: 10.3389/fpsyg.2020.00721

[27] JASP Team. JASP (Version 0.16.2)[Computer software], 2022.

[28] P. Jishtu and M. A. Yadav. Application of augmented reality & virtual reality in architecture and planning: an overview. In *Computer Science & Information Technology (CS & IT)*. AIRCC Publishing Corporation, Dec. 2020. doi: 10.5121/csit.2020.102008

[29] J. W. Kelly, L. A. Cherep, B. Klesel, Z. D. Siegel, and S. George. Comparison of two methods for improving distance perception in virtual reality. *ACM Trans. Appl. Percept.*, 15(2), mar 2018. doi: 10.1145/3165285

[30] S. J. Kim, Y. Jeong, S. Park, K. Ryu, and G. Oh. A survey of drone use for entertainment and avr (augmented and virtual reality). In *Augmented reality and virtual reality*, pp. 339–352. Springer, 2018.

[31] S. Kuliga, T. Thrash, R. Dalton, and C. Hölscher. Virtual reality as an empirical research tool — exploring user experience in a real building and a corresponding virtual model. *Computers, Environment and Urban Systems*, 54:363–375, 2015. doi: 10.1016/j.compenvurbsys.2015.09.006

[32] Lager, Mårten and Topp, Elin Anna and Malec, Jacek. Remote Operation of Unmanned Surface Vessel through Virtual Reality, 2018.

[33] M. Le Chénéchal, T. Duval, V. Gouranton, J. Royan, and B. Arnaldi. Help! i need a remote guide in my mixed reality collaborative environment. *Frontiers in Robotics and AI*, 6, 2019. doi: 10.3389/frobt.2019.00106

[34] C.-L. Li, M. P. Aivar, D. M. Kit, M. H. Tong, and M. M. Hayhoe. Memory and visual search in naturalistic 2d and 3d environments. *Journal of Vision*, 16(8), Jun 2016. doi: 10.1167/16.8.9

[35] M. Magrini, O. Curzio, M. Tampucci, G. Donzelli, L. Cori, M. C. Imiotti, S. Maestro, and D. Moroni. Anorexia nervosa, body image perception and virtual reality therapeutic applications: State of the art and operational proposal. *International Journal of Environmental Research and Public Health*, 19(5):2533, Feb. 2022. doi: 10.3390/ijerph19052533

[36] S. C. Mallam, S. Nazir, and S. K. Renganayagalu. Rethinking maritime education, training, and operations in the digital era: Applications for emerging immersive technologies. *Journal of Marine Science and Engineering*, 7(12), 2019. doi: 10.3390/jmse7120428

[37] L. Morís Fernández and M. A. Vadillo. Flexibility in reaction time analysis: many roads to a false positive? *Royal Society Open Science*, 7(2):190831, 2020. doi: 10.1098/rsos.190831

[38] B. Olk, A. Dinu, D. J. Zielinski, and R. Kopper. Measuring visual search and distraction in immersive virtual reality. *Royal Society Open Science*, 5(5):172331, 2018. doi: 10.1098/rsos.172331

[39] E. Peillard, T. Thebaud, J.-M. Normand, F. Argelaguet, G. Moreau, and A. Lécuyer. Virtual objects look farther on the sides: The anisotropy of distance perception in virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 227–236, 2019. doi: 10.1109/VR.2019.8797826

[40] D. J. Pieramici, F. Heimann, R. Brassard, G. Barteselli, and S. Ranade. Virtual Reality Becomes a Reality for Ophthalmologic Surgical Clinical Trials. *Translational Vision Science & Technology*, 9(7):1–1, 06 2020. doi: 10.1167/tvst.9.7.1

[41] J. Poppenk, S. Köhler, and M. Moscovitch. Revisiting the novelty effect: when familiarity, not novelty, enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(5):1321, 2010.

[42] A. E. Richardson, D. R. Montello, and M. Hegarty. Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition*, 27(4):741–750, Jul 1999. doi: 10.3758/bf03211566

[43] M. Rubo and M. Gamer. Virtual reality as a proxy for real-life social attention? In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, ETRA '18. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3204493.3207411

[44] P.-L. Sanchez-Gonzalez, D. Díaz-Gutiérrez, T. J. Leo, and L. R. Núñez-Rivas. Toward digitalization of maritime transport? *Sensors*, 19(4), 2019. doi: 10.3390/s19040926

[45] K. F. Schulz, D. G. Altman, and D. Moher. Consort 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*, 340, 2010. doi: 10.1136/bmj.c332

[46] H. Sun, R. T. Wright, and J. B. Thatcher. Revisiting the impact of system use on task performance: An exploitative-explorative system use framework. *Journal of the Association for Information Systems*, p. 398–433, Apr 2019. doi: 10.17705/1jais.00539

[47] Mini mouse macro, 2020.

[48] G. R. Van den Bos and A. P. Association. Ceiling effect.

[49] J. van Doorn, D. van den Bergh, U. Böhm, F. Dablander, K. Derks, T. Draws, A. Etz, N. J. Evans, Q. F. Gronau, J. M. Haaf, and et al. The jasp guidelines for conducting and reporting a bayesian analysis. *Psychonomic Bulletin & Review*, 28(3):813–826, Oct 2020. doi: 10.3758/s13423-020-01798-5

[50] C. J. Wilson and A. Soranzo. The use of virtual reality in psychology: A case study in visual perception. *Computational and Mathematical Methods in Medicine*, 2015:1–7, Aug 2015. doi: 10.1155/2015/151702

[51] T. V. Zandt. Analysis of response time distributions. In J. Wixted and H. Pashler, eds., *Stevens' Handbook of Experimental Psychology*. Wiley, 2002.

[52] L. H. Zeuwts, E. Iliano, M. Smith, F. Deconinck, and M. Lenoir. Mental fatigue delays visual search behaviour in young cyclists when negotiating complex traffic situations: A study in virtual reality. *Accident Analysis & Prevention*, 161:106387, 2021. doi: 10.1016/j.aap.2021.106387

[53] P. Zuesongdham. Port of hamburg – virtual reality for model-based port infrastructure management, 2017.