Detecting Spatial Orientation Demands during Virtual Navigation using EEG Brain Sensing

Thinh Nguyen-Vo
School of Interactive Arts and Technology
Simon Fraser University
Canada
tnguyenv@sfu.ca

Steve DiPaola
School of Interactive Arts and Technology
Simon Fraser University
Canada
sdipaola@sfu.ca

Bernhard E. Riecke
School of Interactive Arts and Technology
Simon Fraser University
Canada
ber1@sfu.ca

Figure 1. Left: A participant wearing HTC VIVE HMD and the EMOTIV Insight EEG BCI. Right: Visual stimuli of the virtual environment designed to provide optic flow but eliminate all visual landmarks.

Abstract

This study shows how brain sensing can offer insight to the evaluation of human spatial orientation in virtual reality (VR) and establish a role for electroencephalogram (EEG) in virtual navigation. Research suggests that the evaluation of spatial orientation in VR benefits by going beyond performance measures or questionnaires to measurements of the user's cognitive state. While EEG has emerged as a practical brain sensing technology in cognitive research, spatial orientation tasks often rely on multiple factors (e.g., reference frame used, ability to update simulated rotation, and/or left-right confusion) which may be inaccessible to this measurement. EEG has been shown to correlate with human spatial orientation in previous research. In this paper, we use convolutional neural network (CNN), an advanced technique in machine learning, to train a detection model that can identify moments in which VR users experienced some increase in spatial orientation demands in real-time. Our results demonstrate that we can indeed use machine learning technique to detect such cognitive state of increasing spatial orientation demands in virtual reality research with 96% accurate on average.

CCS Concepts • **Human-centered computing** → **Virtual reality**; • ***Human-centered computing*** → *Systems and tools for interaction design*

Keywords Convolutional Neural Network, Detection Model, EEG, Machine Learning, Spatial Orientation, Virtual Reality

1 Introduction

The quantitative evaluation of psychological outcomes has been a significant goal of both the cognitive science and human-computer interaction (HCI) / VR community for decades. Many quantitative and qualitative approaches have been proposed to peek into the user's cognitive processes during virtual experiences. Nevertheless, there are potential confounds to evaluate psychological outcome without directly monitoring the brain's cognitive processes. Evaluations of basic tasks may have low external validity to the real-life condition. Psychophysical research often uses artificial or abstract stimuli and situations tend to be not appealing to the participants, which might reduce real-world transferability of the results [24, 25]. Moreover, previous psychology research suggests that evaluating performance without also assessing workload may lead to incorrect conclusions about the cognitive efficiency of an interface [13, 27, 41]. Finally, cognitive states can change even when performance remains stable. In other words, performance metrics may not always accurately reflect underlying cognitive processes [7, 39].

Consequently, there have been many approaches proposed in objective methods to evaluate cognitive processes in HCI research [14, 33]. In particular, EEG has been widely used as a practical brain sensing technology in HCI as it is relatively affordable and easy to use. It has been proven to be effective in measuring brain cognitive load [11]. Another study also reports the correlation between EEG and insightful problem solving [35]. In cognitive science, spatial orientation refers to an ability of most animals including human that allows them to navigate through their immediate environment and interacting with it effectively with little cognitive effort. This ability is essential for most animals' daily activities such as finding food and homing. In VR, due to multiple factors such as sensory conflicts and latency, human spatial orientation degrades significantly. Previous research has shown people take significantly longer

Nguyen-Vo, T., DiPaola, S., & Riecke, B. E. (2017). Detecting Spatial Orientation Demands during Virtual Navigation using EEG Brain Sensing. ACM SIGPLAN Workshop on Software for Augmented and Virtual Reality (SAVR 2017) (pp. 1–5). Presented at the ACM SIGPLAN Workshop on Software for Augmented and Virtual Reality (SAVR 2017), Vancouver, Canada: ACM. talk: https://youtu.be/-2gKhvxUGxg

time keeping track of their positions when traveling in VR than they do in the real world [40]. There are a variety of factors have been shown to affect human spatial updating process in virtual environments (e.g., [23, 26, 28]), such as left-right confusion [10, 29] and inability to update visually simulated rotations [17, 40].

In order to measure the effectiveness of spatial orientation, previous studies used multiple experimental tasks that require maintaining spatial awareness in VR such as rapid pointing (e.g., [17, 32]), point-to-origin (e.g., [10, 16, 17, 29]) or navigational search (e.g., [19, 26, 30, 34]). However, these tasks, as discussed above, often have low external validity. On the other hand, EEG or theta-band frequency oscillation has been shown to be correlated with human spatial orientation in multiple contexts (e.g., [4, 9]). For example, striking episodes of high-amplitude theta band oscillations has been revealed during virtual maze navigation [15]. The association between theta oscillation and the performance of navigational tasks in humans is also suggested in another study [1].

In this work, we test the viability of using EEG to observe how spatial orientation demands modifies brain activity in virtual navigation. We conducted two experiments to (1) examine how participants process spatial orientation demand in their brains, and (2) determine the efficacy of using machine learning, more specifically convolutional neural network (CNN), as a technique for detecting spatial orientation demands in virtual navigation.

To investigate this, we first conducted a controlled experiment with a two-level independent variable of spatial cognition: without or with a spatial orientation demand in the form of a point-to-origin task. For the first condition, participants only passively observed a simulated excursion in VR. For the latter condition, in addition to a simulated excursion, participants were asked to do a point-to-origin task that is to keep track of their translation and rotation so that at the end of the trajectory, they can point back to the origin which is the location where they started the navigation. The dependent variable is the power of theta-band oscillation extracted from the collected EEG data. Secondly, we proposed a CNN architecture and used cross-validation to evaluate the efficacy of this network in detecting spatial orientation demand during virtual navigation. Based on our results, we make two contributions:

- Our findings suggest that EEG can be used to monitor differences in brain activity that derive from spatial orientation demands during virtual navigation. We find that the average power of theta-band oscillation differs significantly during navigation with/without spatial orientation demands. That is to say changes in theta band rhythms correlated with the presence of spatial orientation demands.
- We propose that such machine learning technique as CNN can be effectively used to detect spatial orientation demands in human brain activity. The proposed network architecture can effectively detect orientation demand with average accuracy of 96%. This result also suggested that advanced AI techniques can be adopted in analyzing different kind of psychophysical data in order to enhance the objectivity in measuring psychological outcomes in HCI research.

2 Method

2.1 Stimuli and Apparatus

Virtual Environment

Participants wore a head-mounted display (HMD) displaying a virtual world. As depicted in Fig. 1, this virtual environment consists of a star field of randomly distributed white spheres in a black environment, which was designed to provide strong optic flow but no landmarks or any other orientation cue. This optic flow is the only visual cue that participants can use to keep track of their position and rotation during the virtual navigation. There is a large body of literature showing that optic flow can induce a compelling illusion of self- motion (vection) [6, 12, 31, 37]. In addition, previous studies have also shown that humans can extract turning angles and travel distances from pure optic flow information [5, 20, 37, 38].

Visualization

HTC Vive HMD has been used to binocularly present the virtual environment to participants in this experiment. The HTC Vive provides a per-eye resolution of 1080 x 1200 pixel and a binocular FOV of 110° diagonally. Stimuli were generated in real time at 90Hz.

Tracking and Interaction

The head tracking embedded in the HMD was enabled. The wireless controller going with the HMD was also used as a means for pointing to the origin at the end of each trajectory.

Brain Computer Interface (BCI)

EEG data were collected using a five-channel dry electrode headset (EMOTIV Insight). Five electrodes were located at standard coordinates: AF3, AF4, T7, T8, O1. The data measured is oversampled at 2048Hz/channel and this is filtered to remove all traces of environmental electromagnetic interference and then down sampled to 128Hz.

Task

Initially, participants were positioned in the middle of a star field. They had time to look around and get familiar with the virtual world. When participants were ready, they were asked to sit still on a stable chair. After a few seconds, the navigation would begin. Participants were passively navigated through a predefined trajectory including straight and curvy segments. There were five trajectories and their order was randomized for each condition. There is a slight difference between two conditions:

- Normal: In baseline condition, participants were asked to simply watch the navigation.
- Spatial Updating / Orientation Demanding: In the spatial orientation demanding condition, participants were asked to keep track on their position and rotation changes so that at the end of the trajectory, they can point back to the origin of the current navigation using the HTC Vive controller.

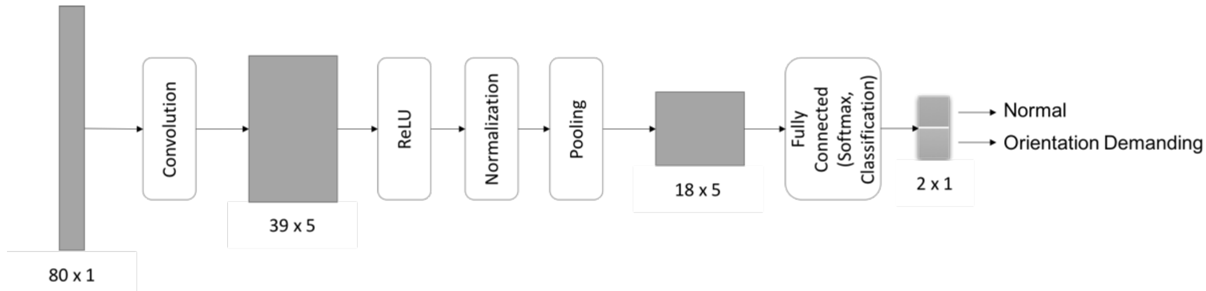

Figure 2. The proposed topology of five-layer CNN

2.2 Procedure

Participants began by reading written instructions explaining the task as well as the use of the equipment. Each participant completed one trial in the training section to get familiar with the interface and five trials for each of the two variations of the task. Each trial lasted approximately 30 seconds. Participants had 15-second breaks on average between trials.

2.3 Experimental Design

This is a repeated-measures experimental design where the independent variable is the variation of the task (2 levels: normal and spatial orientation demanding). Therefore, every participant took part in all two conditions, in counter-balanced order. Oscillatory responses are the most common and well-studied characteristics in EEG [2, 21]. In this study, we hence measured theta-band oscillation power as the dependent variable. Data were then analyzed using repeated-measures ANOVA and used as dataset for training and testing CNN-based detection model.

2.4 Dataset

In this study, the EEG dataset was collected during our own spatial updating experiment using a EMOTIV Insight BCI. The original 20 30-second datasets (2 participants x 2 conditions x 5 trials) was pre-processed to filter out noises and normalize lengths between trials. Pre-processed datasets included twelve 20-second datasets (2 participants x 2 conditions x 3 trials). At this step, for each condition we had 120 records per participants (2 Hz x 20 seconds x 3 trials) and each record included 4 channels (AF3, T7, T8, AF4). Before entering CNN, the data was reconstructed using a window size of 20 (relevant to 10 seconds). Fig. 3 illustrates this step on how data were constructed to 80-element vectors.

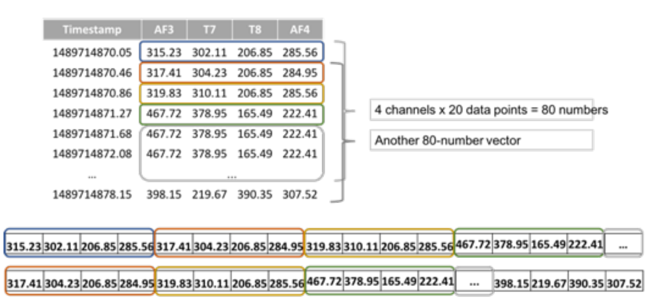

Figure 3. Data preprocessing and formatting using window size of 20

2.5 Convolutional Neural Network

CNN is a neural network which consists of a multilayer perceptron (MLP), and which possesses a special topology containing multiple hidden layers. Our CNN model topology is illustrated in Fig. 2. This CNN has five layers, including convolutional, rectified linear units (ReLU), normalization, max pooling, and fully connected layers in order:

- The input data has a structure of 80 x 1 (as described above). The convolutional layer convolves this input data matrix via five 4 x 1 filters with stride of five and outputs the filtered data map. The structures of these filters were learned by back propagation during the training procedure.
- The filtered data maps whose structure of 39 x 5, are then entered ReLU, cross-channel normalization, and max pooling layers. This max pooling layer used a filter with size of five and stride of two. After this step, the featured map has the size of 18 x 5.
- Finally, the features obtained are transferred to the fully connected layer. Ultimate classification is based on these features. Softmax function is used as the activation function.

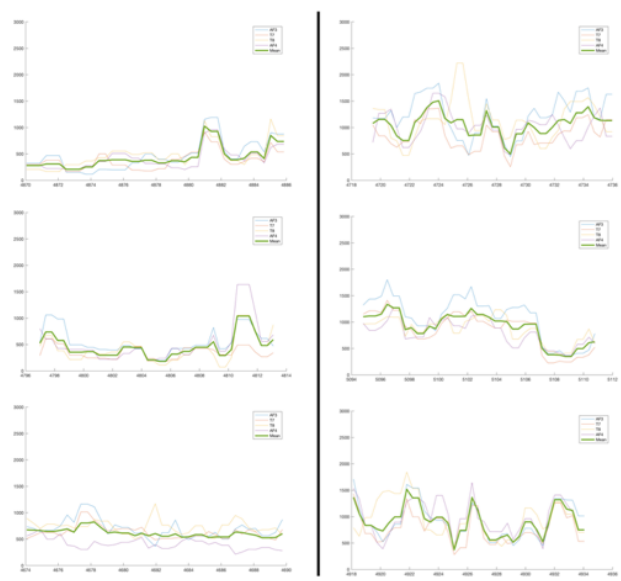

Figure 4. Theta-band oscillation power of one participant in normal (left) and spatial orientation demanding (right) conditions.

3 Results

Visual inspection of the theta-band oscillation revealed striking changing patterns between the two conditions of spatial orientation demands. Fig. 4 shows an example of the differences in theta-band rhythm. While the signals seemed to be more stable in baseline condition, they fluctuated significantly stronger in orientation demanding condition.

Theta-band oscillation power is summarized in Fig. 5 and was analyzed using repeated-measures ANOVA. Our result confirmed previous findings that **spatial orientation demands significantly change theta-band oscillation power during virtual navigation**. ANOVA revealed a significant effect of spatial orientation on EEG signal, $F(1, 463) = 271.44$, $p < .0001$.

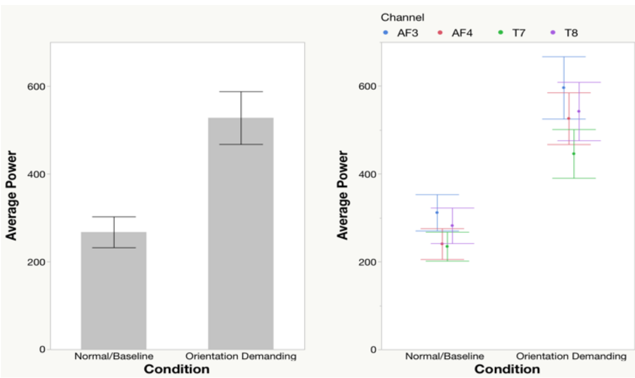

Figure 5. Mean data of theta-band oscillation power (Left: cross-channel; Right: individual channels). Error bars indicate confidence intervals (CI = .95)

Cross-validation on three-fold datasets with two-fold training and one-fold testing showed the efficacy of the proposed CNN with the average accuracy of 96% for inter-subject condition (SD = 8.57%) and 63% for intra-subject condition (SD = 8.76%). Intra-subject is the condition where the model was built on data from a single participant (using 2 folds), then used the trained model to test on himself/herself (using the remaining fold). The inter-subject is where we use data from all participants together to train and test. **This result suggests that CNN is robust for spatial orientation demands detection in inter-subject condition**.

4 Discussion and Conclusion

There is a large body of literature demonstrating the relationship between theta oscillation and brain processes underlying mental workload, working memory [18, 22, 36] or spatial orientation [1, 15]. The current study confirmed previous findings about the correlation between theta oscillation and human spatial orientation in the context of a point-to-origin task in VR. We have demonstrated that EEG is a viable technology for investigating the impact of spatial orientation demand on a person's cognition processes. Using the statistical comparison, we found that increasing powers of theta oscillation correlated with the virtual navigation that requires additional spatial orientation demands.

The proposed method of applying machine learning, based on orientation state EEG using CNN, may represent an appropriate technique to develop EEG-based biometric systems which supply good detection performance. This form of biometrics, in addition to providing a higher level of objective measure in spatial orientation research, would have other useful applications, e.g. evaluation metric or interactive inputs for virtual reality application.

The high accuracy for intra-subject (96%) and low (63%) for inter-subject condition also confirmed previous findings about the differences in EEG between individuals. Berhout and Walter observed that EEG might provide different signals between individuals, as pertains to the anatomical and functional traits of their brains [3]. Furthermore, an individual's EEG has been shown to be both stable and specific [8]. In other words, EEG tend to yield small intra-personal differentiation and large inter-personal differentiation, which is why it is easier to build custom detection model for each person than to build a general detection model that can be used for everyone.

In the future, this CNN-based system should be tested on a larger population, with different network topologies and configurations, providing further confirmation of the application of the system as well as the robustness of machine learning techniques on EEG analysis. We believe that EEG will become more practical in the near future and will likely be embedded in future HMDs, and that recognizing such cognitive states as spatial cognition or motion using commercial-grade BCI will become much more widespread. For that reason, this is an open direction for future research to continue exploring the potential of EEG for cognitive state detection and recognition.

References

[1] Araújo, D.B. de et al. 2002. Theta Oscillations and Human Navigation: A Magnetoencephalography Study. *Journal of Cognitive Neuroscience.* 14, 1 (Jan. 2002), 70–78. DOI:https://doi.org/10.1162/089892902317205339.

[2] Bashivan, P. et al. 2014. Spectrotemporal dynamics of the EEG during working memory encoding and maintenance predicts individual behavioral capacity. *European Journal of Neuroscience.* 40, 12 (Dec. 2014), 3774–3784. DOI:https://doi.org/10.1111/ejn.12749.

[3] Berkhout, J. and Walter, D.O. 1968. Temporal Stability and Individual Differences in the Human EEG: An Analysis of Variance of Spectral Values. *IEEE Transactions on Biomedical Engineering.* BME-15, 3 (Jul. 1968), 165–168. DOI:https://doi.org/10.1109/TBME.1968.4502560.

[4] Bischof, W.F. and Boulanger, P. 2003. Spatial Navigation in Virtual Reality Environments: An EEG Analysis. *CyberPsychology & Behavior.* 6, 5 (Oct. 2003), 487–495. DOI:https://doi.org/10.1089/109493103769710514.

[5] Bremmer, F. and Lappe, M. 1999. The use of optical velocities for distance discrimination and reproduction during visually simulated self motion. *Experimental Brain Research.* 127, 1 (Jun. 1999), 33–42. DOI:https://doi.org/10.1007/s002210050771.

[6] Dichgans, J. and Brandt, T. 1978. Visual-Vestibular Interaction: Effects on Self-Motion Perception and Postural Control. *Perception.* Springer. 756–804.

[7] Dick Waard, D. 1996. The measurement of drivers' mental workload. (1996).

[8] van Dis, H. et al. 1979. Individual differences in the human electroencephalogram during quiet wakefulness. *Electroencephalography and Clinical Neurophysiology.* 47, 1 (Jul. 1979), 87–94.

[9] Ekstrom, A.D. et al. 2005. Human hippocampal theta activity during virtual navigation. *Hippocampus.* 15, 7 (Jan. 2005), 881–889. DOI:https://doi.org/10.1002/hipo.20109.

[10] Gramann, K. et al. 2005. Evidence of Separable Spatial Representations in a Virtual Navigation task. *Journal of Experimental Psychology – Human Perception and Performance.* 31, 6 (Dec. 2005), 1199–1223.

[11] Grimes, D. et al. 2008. Feasibility and Pragmatics of Classifying Working Memory Load with an Electroencephalograph. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2008), 835–844.

[12] Hettinger, L.J. et al. 2014. Illusory Self-motion in Virtual Environments. *Handbook of Virtual Environments.* K.S. Hale and K.M. Stanney, eds. CRC Press. 435–466.

[13] Huang, W. et al. 2009. Measuring Effectiveness of Graph Visualizations: A Cognitive Load Perspective. *Information Visualization.* 8, 3 (Jun. 2009), 139–152. DOI:https://doi.org/10.1057/ivs.2009.10.

[14] Huang, W. 2007. Using eye tracking to investigate graph layout effects. *2007 6th International Asia-Pacific Symposium on Visualization* (Feb. 2007), 97–100.

[15] Kahana, M.J. et al. 1999. Human theta oscillations exhibit task dependence during virtual maze navigation. *Nature.* 399, 6738 (Jun. 1999), 781–784. DOI:https://doi.org/10.1038/21645.

[16] Kitson, A. et al. 2016. Influence of Ethnicity, Gender and Answering Mode on a Virtual Point-to-Origin Task. *Frontiers in Behavioral Neuroscience.* 10, (Feb. 2016). DOI:https://doi.org/10.3389/fnbeh.2016.00022.

[17] Klatzky, R.L. et al. 1998. Spatial Updating of Self-Position and Orientation During Real, Imagined, and Virtual Locomotion. *Psychological Science.* 9, 4 (1998), 293–298. DOI:https://doi.org/10.1111/1467-9280.00058.

[18] Klimesch, W. 1996. Memory processes, brain oscillations and EEG synchronization. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology.* 24, 1–2 (Nov. 1996), 61–100.

[19] Lessels, S. and Ruddle, R.A. 2005. Movement Around Real and Virtual Cluttered Environments. *Presence: Teleoperators and Virtual Environments.* 14, 5 (Oct. 2005), 580–596. DOI:https://doi.org/10.1162/105474605774918778.

[20] Loomis, J.M. et al. 1999. Immersive virtual environment technology as a basic research tool in psychology. *Behav. Res. Methods Instr. Comput.* 31, 4 (Nov. 1999), 557–564.

[21] Lotte, F. et al. 2007. A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering.* 4, 2 (Jun. 2007), R1–R13. DOI:https://doi.org/10.1088/1741-2560/4/2/R01.

[22] Maguire, E.A. et al. 1999. Human spatial navigation: cognitive maps, sexual dimorphism, and neural substrates. *Current Opinion in Neurobiology.* 9, 2 (Apr. 1999), 171–177.

[23] Moura, D. and Riecke, B.E. 2009. Is Seeing a Virtual Environment Like Seeing the Real Thing? *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization* (New York, NY, USA, 2009), 131–131.

[24] Neisser, U. 1976. *Cognition and reality: Principles and implications of cognitive psychology.* W H Freeman/Times Books/ Henry Holt & Co.

[25] Newell, A. 1973. *You can't play 20 questions with nature and win: Projective comments on the papers of this symposium.* Carnegie Mellon University.

[26] Nguyen-Vo, T. et al. 2017. Moving in a box: Improving spatial orientation in virtual reality using simulated reference frames. *2017 IEEE Symposium on 3D User Interfaces (3DUI)* (Mar. 2017), 207–208.

[27] Paas, F.G.W.C. and Van Merriënboer, J.J.G. 1993. The Efficiency of Instructional Conditions: An Approach to Combine Mental Effort and Performance Measures. *Human Factors.* 35, 4 (Dec. 1993), 737–743. DOI:https://doi.org/10.1177/001872089303500412.

[28] Richardson, A.E. et al. 1999. Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition.* 27, 4 (Jul. 1999), 741–750. DOI:https://doi.org/10.3758/BF03211566.

[29] Riecke, B.E. 2008. Consistent Left-Right Reversals for Visual Path Integration in Virtual Reality: More than a Failure to Update One's Heading? *Presence: Teleoperators and Virtual Environments.* 17, 2 (2008), 143–175. DOI:https://doi.org/10.1162/pres.17.2.143</p>.

[30] Riecke, B.E. et al. 2010. Do we need to walk for effective virtual reality navigation? physical rotations alone may suffice. *Spatial Cognition VII.* Springer. 234–247.

[31] Riecke, B.E. 2006. Simple user-generated motion cueing can enhance self-motion perception (Vection) in virtual reality. *Proceedings of the ACM symposium on Virtual reality software and technology* (Limassol, Cyprus, 2006), 104–107.

[32] Riecke, B.E. et al. 2007. Spatial updating in virtual reality: The sufficiency of visual information. *Psychological Research.* 71, 3 (May 2007), 298–313. DOI:https://doi.org/10.1007/s00426-006-0085-z.

[33] Rowe, D.W. et al. 1998. Heart Rate Variability: Indicator of User State As an Aid to Human-computer Interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 1998), 480–487.

[34] Ruddle, R.A. and Lessels, S. 2006. For efficient navigational search, humans require full physical movement, but not a rich visual scene. *Psychological Science.* 17, 6 (Jun. 2006), 460–465. DOI:https://doi.org/10.1111/j.1467-9280.2006.01728.x.

[35] Sandkühler, S. and Bhattacharya, J. 2008. Deconstructing insight: EEG correlates of insightful problem solving. *PloS One.* 3, 1 (Jan. 2008), e1459. DOI:https://doi.org/10.1371/journal.pone.0001459.

[36] Schack, B. et al. 2005. Phase synchronization between theta and upper alpha oscillations in a working memory task. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology.* 57, 2 (Aug. 2005), 105–114. DOI:https://doi.org/10.1016/j.ijpsycho.2005.03.016.

[37] Warren, R. and Wertheim, A.H. 1990. *Perception & Control of Self-motion.*

[38] Warren, W.H. et al. 2001. Optic flow is used to control human walking. *Nature Neuroscience.* 4, 2 (Feb. 2001), 213–216. DOI:https://doi.org/10.1038/84054.

[39] Wickens, C.D. et al. 2015. *Engineering Psychology & Human Performance.* Psychology Press.

[40] Witmer, B.G. et al. 1996. Virtual spaces and real world places: transfer of route knowledge. *International Journal of Human-Computer Studies.* 45, 4 (1996), 413–428.

[41] Yeh, Y.-Y. and Wickens, C.D. 1988. Dissociation of Performance and Subjective Measures of Workload. *Human Factors.* 30, 1 (Feb. 1988), 111–120. DOI:https://doi.org/10.1177/001872088803000110.