Virtual Reality (2024) 28:174
https://doi.org/10.1007/s10055-024-01067-z

ORIGINAL ARTICLE

Physiology-driven cybersickness detection in virtual reality: a machine learning and explainable AI approach

Javad Sameri[1] · Hendrick Coenegracht[1] · Sam Van Damme[1,2] · Filip De Turck[1] · Maria Torres Vega[1,2]

Received: 27 October 2023 / Accepted: 29 October 2024

Abstract

One of the major obstacles to the widespread adoption of Virtual Reality (VR) is Cybersickness. It is a sense of physical discomfort akin to motion sickness experienced by the users either during or subsequent to VR utilization. Typically, it is detected through explicit methods such as self-reported questionnaires, which are not ideal for online and implicit monitoring of users' well-being. This study tackles the challenge of implicitly detecting and measuring Cybersickness in VR through physiological signals. Therefore, we propose a multimodal approach that integrates physiological signals with self-reported measures. We utilize a mixed-methods design combining quantitative and qualitative analyses, using a roller coaster simulation as the experimental paradigm. The research analyzes physiological data collected from a group of 22 participants exposed to a simulated VR rollercoaster experience, through statistical analysis and machine learning algorithms. The physiological markers studied include Electroencephalograms (EEG), Electrodermal Activity, Blood Volume Pulse, and skin Temperature signals. Additionally, the study elucidates the complex interplay among physiological markers using Explainable AI (XAI). We found out significant correlations between high Simulator Sickness Questionnaire scores and physiological measures such as brain rhythms and EEG indices related to engagement, visual fatigue, and drowsiness, as well as heart rate variability. Moreover, the proposed machine learning model achieves an accuracy of 86.66 % in detecting elevated Cybersickness symptoms, which is higher than the accuracy achieved by existing techniques. Further exploration using the XAI technique confirms the elevated levels of drowsiness and reduced engagement in participants experiencing elevated Cybersickness. By providing a comprehensive, multimodal approach for quantitative assessment, this study fills a gap in the existing literature and paves the path for the development of adaptive systems that modulate their behavior based on physiological data.

Keywords Cybersickness · Virtual reality · Physiological signals · Explainable AI · Brain–computer interface

1 Introduction

The advent of (XR) has revolutionized various sectors, from healthcare and education to entertainment and design, offering unprecedented immersive experiences (Gans and Nagaraj 2023). However, the widespread adoption of immersive technologies has given rise to a new set of challenges. One of the most fundamental ones is Cybersickness (Yildirim 2020).

This phenomenon is marked by physical discomfort closely resembling motion sickness and is triggered by non-comfortable motions as experienced by the users in the virtual environment. Cybersickness manifests itself in a significant portion of XR users, between 60–80% depending on the experimental setup, through various symptoms such

✉ Javad Sameri
Javad.Sameri@UGent.be

Hendrick Coenegracht
Hendrik.Coenegracht@UGent.be

Sam Van Damme
Sam.VanDamme@UGent.be

Filip De Turck
Filip.DeTurck@UGent.be

Maria Torres Vega
Maria.TorresVega@UGent.be

[1] IDLab, Department of Information Technology (INTEC), Ghent University- imec, Ghent, Belgium

[2] eMedia Research Lab, Department of Electrical Engineering (ESAT), KU Leuven, Leuven, Belgium

Published online: 19 November 2024

as nausea, dizziness, and headache (Tian et al. 2022; Garrido et al. 2022; Chang et al. 2023).

In order to effectively address Cybersickness, it is important to be able to accurately detect it (Fernandes and Feiner 2016; Maeda et al. 2005), which is a very complex task. Several methods have been used to identify and quantify Cybersickness, including self-reported measures such as the Simulator Sickness Questionnaire (SSQ), as well as behavioral tracking methods, e.g., head movements and gaze patterns (Teixeira and Palmisano 2021; Clifton and Palmisano 2020; Fernandes and Feiner 2016).

Physiological signals present a compelling subjectively implicit, fast evaluation of Cybersickness. Metrics such as heart rate variability, skin conductance, and electroencephalographic patterns serve as robust indicators of an individual's physiological state (Chang et al. 2023; Islam et al. 2021). Some studies have also found a significant correlation between brain rhythms and Cybersickness, suggesting that investigating brain rhythms and other physiological responses in Cybersickness-inducing environments could lead to further insights (Krokos and Varshney 2022; Chang et al. 2023; Kogler et al. 2021). Unfortunately, these results are very limited and tailored to specific applications.

This paper seeks to understand the onset and severity of Cybersickness in (VR) environments by correlating physiological markers with self-reported symptoms. We utilize a mixed-methods research design, incorporating both quantitative and qualitative analyses. Specifically, we employ statistical analysis and machine learning algorithms to examine the physiological data from our experiment. In detail, this paper makes four main contributions:

1. A VR testbed is established for the experimental evaluation of Cybersickness. This includes a VR roller coaster, specifically developed for this study, physiological logging, subjective questionnaires, and a dataset of 22 user data anonymized and compliant with the General Data Protection Regulation (GDPR). This dataset has been made publicly available[1].
2. A comprehensive correlation analysis is conducted to explore the relationship between physiology and the subjective perception of Cybersickness. The findings demonstrate significant correlations between elevated SSQ scores and physiological markers such as brain rhythms, and heart rate fluctuations.
3. A predictive methodology is developed for Cybersickness detection, drawing upon the physiological markers identified.
4. An Explainable AI (XAI) technique is employed to elucidate the physiological correlations pinpointed by the predictive model. For instance, it reveals Cybersickness proneness of subjects with elevated levels of fatigue and drowsiness.The remainder of this paper is organized as follows. Section 2 provides a review of related literature studies. Section 3 explains the experimental methodology and data acquisition procedures for the Cybersickness framework. Subsequently, Sect. 4 delves into an analysis of the collected data. In Sect. 5, we provide details on the construction of the physiology-driven Cybersickness model. Section 6 discusses the model's interpretability using an XAI. We present our conclusions in Sect.7, where the potential of Cybersickness physiological modeling and future research directions are delineated.

2 Related work

Many studies have aimed to understand the mechanisms and causes of Cybersickness, but a clear theory is still lacking. Three main theories are often discussed: (i) Sensory Conflict Theory, (ii) Intoxication Theory, and (iii) Postural Instability Theory (Qu et al. 2022).

(i) The Sensory Conflict Theory consists of the central hypothesis that motion sickness is caused by conflict between the current pattern of sensory inputs about self-movement and the pattern that is expected on the basis of previous experience. Furthermore, it also assumes that the degree of motion sickness is proportional to the magnitude of sensory conflict (Warwick-Evans et al. 1998; Reason 1978). However, this theory does not always explain why people react differently to the same conditions.

(ii) Intoxication Theory suggests that the body interprets these mixed signals as a sign of poisoning, leading to multiple symptoms such as vomitting or nausea (Treisman 1977). As such, it bases the explanations for the occurrence of cybersickness in human evolution.

(iii) The postural instability theory, at last, hypotheses that motion sickness is caused by loss of postural control. Furthermore, the degree of motion sickness is assumed to be proportional to the amount of postural instability, which can be manipulated by physical restraint (Warwick-Evans et al. 1998). This challenges Sensory Conflict Theory by focusing on the importance of maintaining balance. Riccio and Stoffregen argue that balance depends on adapting to changes in the environment, such as moving from solid to slippery ground, and that this can increase stress when balance is threatened (Qu et al. 2022; Riccio and Stoffregen 1991; Tian et al.

[1] https://bit.ly/48A9fPp.

2022).For a more detailed description of the aforementioned theories and the concept of cybersickness in general, the reader is referred to the works of Rebenitsch and Owen (2016), Kennedy et al. (2010), Keshavarz and Golding (2022).

Prior research indicated that the severity of Cybersickness tends to be greater in VR environments compared to traditional screen-based simulators, largely attributed to the broader virtual field of view in VR (Stanney et al. 1997; Sharples et al. 2008). Since a positive correlation between a user's virtual field of view and their SSQ score has been established, some mitigation strategies focus on modifying the virtual field of view (Fernandes and Feiner 2016). Innovative solutions have also been explored, such as using galvanic vestibular stimulation to simulate a sense of movement (Maeda et al. 2005; Cevette et al. 2012; Groth et al. 2022). Another approach enhanced the user's perception of self-motion through auditory cues and tactile feedback. From a network perspective, some researchers aim to provide novel approaches to keep motion-to-photon latency consistently below 20 ms to reduce sensory discordance and mitigate Cybersickness (Torres Vega et al. 2020). However, effective mitigation first necessitates reliable metrics for quantifying and differentiating Cybersickness.

The most common method for assessing Cybersickness involves the use of self-report questionnaires, with the SSQ being a standard instrument (Kennedy et al. 1993). The SSQ evaluates sixteen symptoms, categorized into three domains: oculomotor disturbances, disorientation, and nausea. While these self-report measures offer valuable insights, they come with limitations. Administering the questionnaire post-simulation sacrifices the temporal resolution of symptom tracking. Alternatively, pausing an ongoing experiment for questionnaire completion can be impractical and may even alter the participant's physiological state, thereby affecting the accuracy of their responses.

To address these constraints, numerous investigations have turned to physiological markers as a means to assess user condition and identify Cybersickness (Krokos and Varshney 2022; Yang et al. 2023; Nam et al. 2022; Garcia-Agundez et al. 2019b). Among these Electroencephalograms (EEG), Electrocardiogram (ECG), Electrodermal Activity (EDA), Breathing Signal, and Electromyography (EMG) offer the advantage of non-invasive, real-time data collection. Specifically, EEG and ECG have garnered considerable attention for their ability to shed light on the cognitive and emotional states of users (Keshavarz et al. 2022; Garcia-Agundez et al. 2019b; Islam et al. 2021).

Various EEG rhythms, specifically delta (0–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (13–24 Hz), and gamma (24–40 Hz), have been empirically linked to cognitive and emotional states as well as to distinct brain functions localized in different lobes (Ismail and Karwowski 2020). Variability in these rhythms, whether observed globally or in targeted regions like the frontal, parietal, and occipital lobes, can signify alterations in mental states (Leuthardt et al. 2004). For instance, elevated levels of alpha and theta activity in the frontal lobe correlate with increased fatigue. Similarly, task difficulty has been shown to reduce parietal alpha activity while elevating frontal theta activity (Borghini et al. 2014; Ryu and Myung 2005; Ismail and Karwowski 2020). Additionally, gamma frequency activities have been found to diminish in drowsy states compared to alert states (Zheng and Lu 2017).

A considerable body of research is dedicated to identifying EEG markers that can be correlated with various mental states. For instance, the Task Load Index (TLI) serves as an indicator of task load, engagement, and attention. TLI is calculated as the ratio of mean frontal midline theta (θ) energy to mean parietal alpha (α) energy and has been observed to increase during diverse cognitive tasks (Poussot-Vassal et al. 2017). Another such marker, the Engagement Index (EI), is designed to assess user engagement and is calculated as the ratio of beta (β) power to the sum of theta (θ) and alpha (α) powers $\frac{\beta}{\theta+\alpha}$. Markers for visual fatigue include an increase in the ratios $\frac{\theta+\alpha}{\beta+\alpha}$ and $\frac{\theta+\alpha}{\beta}$, accompanied by a decrease in the EI (Awais et al. 2017). Additionally, a notable increase in the ratio $\frac{\theta+\alpha}{\beta}$ and decreases in the ratios $\frac{\theta}{\alpha}$ and $\frac{\beta}{\alpha}$ have been associated with fatigue and drowsiness (Charbonnier et al. 2016; Zhao et al. 2012). Drowsiness is a shift from alertness to sleep, causing decreased concentration (Asl et al. 2022). Alpha spindles in EEG, characterized by brief narrowband bursts in the alpha band, have been linked to driver fatigue and serve as an objective measure for its assessment (Simon et al. 2011). An increase in alpha spindle activity in the parietal and occipital regions has also been correlated with distraction during extended tasks (Tietze 2001; Ismail and Karwowski 2020). These EEG rhythms and markers offer promising avenues for further research into the quantification and identification of Cybersickness, an area yet to be exhaustively investigated.

In recent studies, various physiological signals have been employed to quantify and predict Cybersickness (Ahn et al. 2020). Krokos et al. demonstrated the utility of EEG signals, particularly in delta, theta, and alpha frequency bands, for correlating with self-reported Cybersickness (Krokos and Varshney 2022). Concurrently, Yang et al. and Nam et al. employed Functional Connectivity (FC) analysis to identify neural hubs and cognitive disruptions linked to Cybersickness (Yang et al. 2023; Nam et al. 2022). Beyond EEG, Heart Rate (HR) and EDA have also been explored for their potential to provide useful information regarding Cybersickness. Heart Rate Variability (HRV),, arising from heart-brain

interactions and the non-linear dynamics of the Autonomic Nervous System (ANS), has emerged as a promising index for neurocardiac function in the context of Cybersickness (Garcia-Agundez et al. 2019b; Wibirama et al. 2018). The role of EDA remains contentious, with conflicting reports on its correlation with Cybersickness (Guna et al. 2019; Dennison et al. 2016). Li et al. (2022) investigate whether and how the newly proposed Human Vestibular Network (HVN) is involved in flagship consumer VR-induced Cybersickness by simultaneously recording autonomic physiological signals as well as neural signals generated in sensorimotor and cognitive domains. To this end, they apply a multiple regression method to put those cognitive indicators in the context of hvn-based biomarkers to investigate their correlations with Cybersickness ratings. They conclude that moderate cybersickness reduces cognitive control ability and is associated with a suppressed training effect. Tauscher et al. (2020) investigated neural and physiological correlates of Cybersickness in a fully immersive dome projection system. Results showed potential for physiological measures as early indicators of Cybersickness. These collective efforts underscore the multifaceted physiological underpinnings of Cybersickness and the ongoing debate on the most effective markers for its quantification and prediction.

In the field of Cybersickness modeling, several studies have employed physiological signals and machine-learning algorithms to predict symptom severity. Dennison et al. and Keshavarz et al. utilized measures such as Breathing Rate (BR) and EDA, yet emphasized the necessity for further model optimization to enhance accuracy (Dennison et al. 2016; Keshavarz et al. 2022). Liao et al. achieved a remarkable 82.83% classification accuracy using deep learning on EEG data, although the 5-minute recording duration constrains its utility for real-time applications (Liao et al. 2020). Another study reached an 83% accuracy rate using a multi-modal biosignal approach but also called for additional research to improve binary classification outcomes (Recenti et al. 2021). A similar study achieved a peak accuracy of 82% in binary classification, further supporting the potential and limitations of physiological markers in this domain (Garcia-Agundez et al. 2019a). However, these studies generally lack transparency in their modeling mechanisms, rendering the machine-learning models as 'black boxes' that impede further advancements in accurate model development (Doran et al. 2017). Collectively, these studies underscore the promising yet challenging landscape of employing physiological markers for accurate Cybersickness classification.

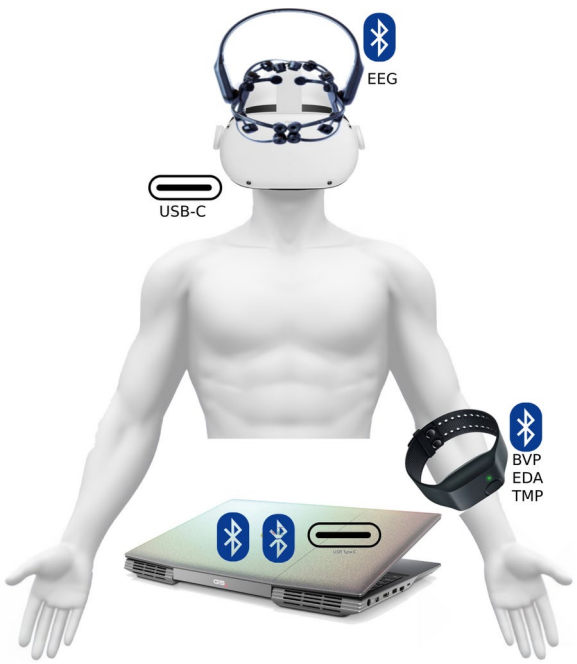


Fig. 1 Schematic representation of the experimental apparatus and configuration

In conclusion, the existing literature highlights a noteworthy correlation between physiological signals and Cybersickness, positioning these signals as viable candidates for non-invasive and implicit evaluations. However, the literature also reveals a significant gap in the performance of multimodal physiological studies, particularly in the quantitative assessment of Cybersickness. This gap underscores the need for a more comprehensive approach that integrates various types of data, a methodological consideration that has yet to be fully addressed in the field.

3 Experimental study of cybersickness

In this section, we establish the groundwork for our experimental study, providing the foundation for further analysis and correlation between explicit subjective perception data and implicit physiological activity data. In Sect. 3.1, we elaborate on the equipment and experimental setup. Section 3.2 outlines the procedure for acquiring the explicit and implicit data from participants. Finally, in Sect. 3.3, we provide a comprehensive explanation of our devised approach for preprocessing and extracting features from physiological recordings.

3.1 Experimental setup

An overview of the employed setup is depicted in Fig. 1. The seamless operation of VR applications necessitates the integration of multiple hardware components. Central to this configuration is the computing unit, where all computational tasks and graphical rendering are executed using Unity 2021. Specifically, a Dell G5 15 gaming laptop equipped with an Intel Core i7-10750 H CPU and an

NVIDIA GeForce RTX 2070 Max-Q Design Graphics Processing Unit (GPU) serves as the computational backbone. For immersive user engagement in the virtual environment, a Meta Quest 2 (HMD)[2] is employed, connected to the computer via a Meta Link[3] cable over a USB 3.0 interface.

Both brain and physiological body signals are collected using the Emotiv EPOC EEG device and the Empatica E4 wristband, respectively (Emotiv 2024; Empatica 2024). Both devices are interfaced with the gaming laptop via Bluetooth connectivity.

The Emotiv EPOC is an EEG device that features metallic electrodes fixed to a plastic base positioned on the scalp to capture brain activity. To optimize signal quality, electrodes are moistened with a saline conductive solution, enhancing conductivity and overall performance.

The headset incorporates a total of 14 electrodes for brain activity measurement, accompanied by two additional electrodes that function as reference points. These electrodes use the international 10–20 system, with positions denoted as AF3, F7, F3, FC5, T7, P7, 01, 02, P8, T8, FC6, F4, F8, and AF4 (Oostenveld and Praamstra 2001).

The Emotive EEG layout is shown in Fig. 2. Raw brain signals across different channels are acquired using the EmotivPro application.[4] The signal is initially sampled at 256 Hz and subsequently resampled to 128 Hz to reduce dimensionality.

Further physiological data is captured utilizing the Empatica E4 wristband. This wearable device facilitates the recording of participants' Blood Volume Pulse (BVP), HR, skin Temperature (TMP), and EDA during the experiment.

Sampling rates vary for distinct sensor measurements: Photoplethysmography (PPG) Sensor at 64 Hz for BVP, Infrared Thermopile at 4 Hz for TMP, EDA sensor at 4 Hz which is considered sufficient for capturing phasic activity (Ronca et al. 2023). The employment of these sensors enables comprehensive data acquisition, allowing a multimodal exploration of both brain and body responses in the context of Cybersickness.

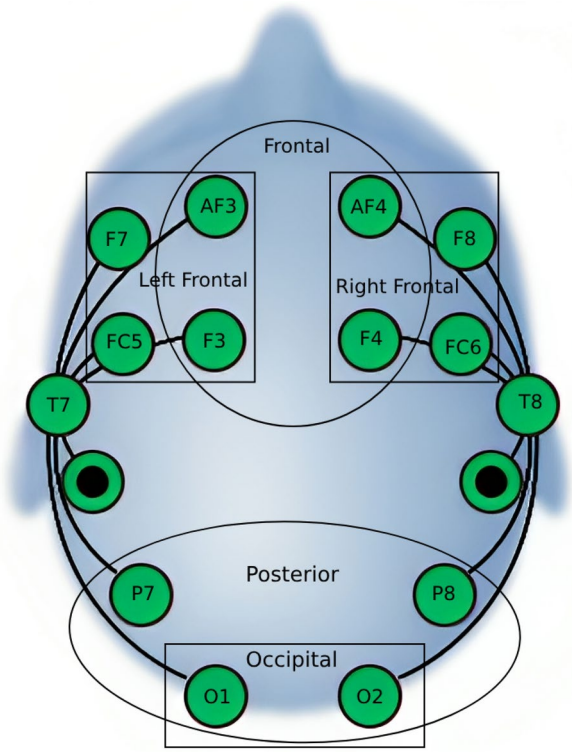


Fig. 2 Location of EEG electrode positions on the scalp and corresponding brain regions

3.2 Subjective evaluation procedure and virtual environment

A user study is conducted to gather subjective perceptions and relevant physiological responses. The experimental protocol starts with a one-minute baseline phase, during which participants are instructed to maintain a motionless posture with closed eyes for the initial 30 s, followed by an additional 30 s with open eyes. This is because physiological signals and mental state vary depending on prior mental and body state. Recording a baseline in EEG and other physiological signals prior to an experiment serves as a control measure to account for individual differences in brain, heart, and other physiological activities. It provides a reference point for normalizing signal changes (Wolpaw et al. 1991; Cohen 2014). A diagram providing an overview of the experiment is shown in Fig. 3.

[2] https://www.oculus.com/experiences/quest/.

[3] https://www.meta.com/be/en/quest/accessories/link-cable/.

[4] https://www.emotiv.com/emotivpro/.

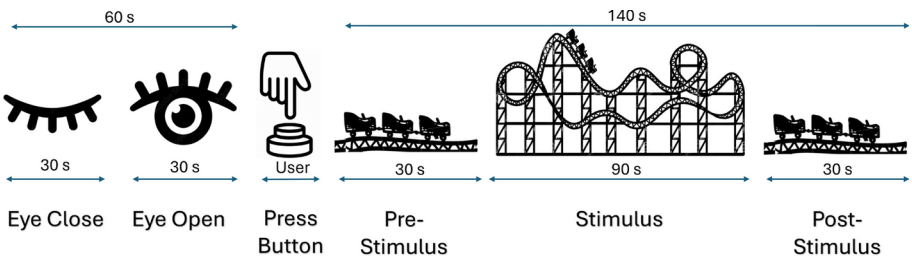


Fig. 3 Schematic of the experimental protocol

(a) A participant engaged in virtual roller coaster experience

(b) Initial scene displayed to participants before the roller coaster simulation.

(c) Comprehensive overview of the roller coaster's trajectory.

Fig. 4 Selected snapshots from the simulation experience

Subsequently, participants are immersed in a virtual environment where they are seated in a VR simulated roller coaster chair, specifically developed by the authors for the purpose of this study. Furthermore, participants were asked to remain seated, with their hands resting on the table or the chair during the experiment, to maintain data integrity and minimize excessive motion artifacts. Figure 4 illustrates a participant engaged with VR roller coaster and two scenes from this simulation. The roller coaster simulation is initiated by the press of the button on Empatica E4 (Fig. 4b) and lasts for a total of 140 s. The simulation is structured into three primary phases. Initially, the roller coaster moves at a mild pace with minimal motion complexity to establish a baseline experience. Subsequently, the speed and motion complexity increase to explore the participants' physiological responses. The final phase sees a significant reduction in both speed and motion complexity. For ease of reference, these phases are termed pre-stimulus, stimulus, and post-stimulus, respectively. Upon completion of the simulation, the HMD, EEG headset, and Empatica device are removed from the participants.

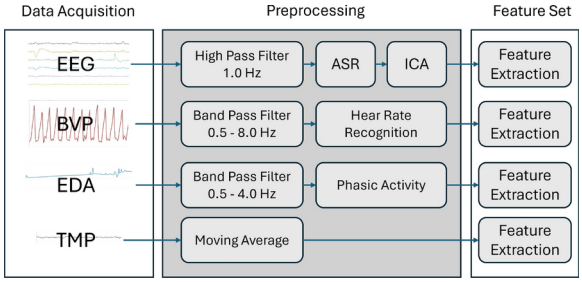


Fig. 5 Workflow for signal preprocessing and feature extraction across physiological modalities

Upon conclusion of the simulation, the participants are instructed to fill out the SSQ as part of the experimental procedure (Kennedy et al. 1993). This post-stimulus assessment evaluates acute symptomatic responses and immediate effects of VR exposure, facilitating further comparison and modeling (Kim et al. 2005; Dennison et al. 2016). Additionally, conducting assessments post-stimulus alleviates concerns related to participants priming (Young et al. 2006). The SSQ was selected for its widespread use and well-established efficacy in assessing Cybersickness, providing consistency with existing research and enabling cross-study comparisons. Despite its limitations, such as potential variability in individual symptom perception, the SSQ remains a valuable tool for quantitatively assessing the immediate effects of VR exposure(Hirzle et al. 2021).

The SSQ comprises 16 questions covering nausea, oculomotor symptoms, and disorientation criteria. These questions are designed to quantify the severity of potential symptoms associated with Cybersickness, allowing for a comprehensive assessment of participants' subjective experiences during the VR roller coaster simulation.

3.3 Physiological data processing

In this section, we outline the procedures for processing and extracting features across each modality. The workflow for each modality involves two primary steps: signal preprocessing to eliminate extraneous information and enhance the quality of the signal, followed by feature extraction to distill meaningful information from the refined signals. It is noteworthy to mention that the approach employed in this paper is executed offline. However, owing to the efficiency of the algorithms, it can be effortlessly implemented in an online paradigm. Finally, the specific processing steps for each modality are illustrated in Fig. 5.

3.3.1 Electroencephalograms (EEG)

EEG signals are renowned for their inherent challenge of low Signal to Noise ratio (SNR) ratio and carrying undesirable information. As a result, signal refinement becomes an integral part of EEG signal processing (Delorme 2023). Consequently, a high-pass filter with a cutoff frequency of 1 Hz is initially employed to remove signal drift caused by skin conductance (Delorme 2023). We employed the EEGLab 'clear raw data' functionality to apply the Adaptive Subspace Reconstruction (ASR) algorithm. This sophisticated implementation of ASR, utilizes Principal Component Analysis (PCA) to identify and eliminate components dominated by artifacts. This identification is conducted by comparing the EEG data to a baseline period deemed free of noise. For this study, the baseline was defined during the initial phase where participants were instructed to remain completely still. The ASR algorithm requires a manually set threshold to determine the acceptable deviation from the clean data. We opted for a lenient threshold of 40, aiming to preserve data integrity while systematically removing conspicuous artifacts (Delorme and Makeig 2004; Delorme 2023; Plechawska-Wojcik et al. 2019). To further enhance signal quality, the Component Analysis (ICA) algorithm is employed to identify and remove ocular artifacts arising from blinks and eye movements (Cohen 2014). EEG processing step has been depicted in Fig. 5.

Following the preprocessing phase, we extracted a series of EEG features for in-depth analysis. Specifically, both relative and absolute Power Spectral Density (PSD) values were computed within the theta ($\theta : 4 - 7_{HZ}$), alpha ($\alpha : 8 - 13_{HZ}$), beta ($\beta : 14 - 29_{HZ}$), and gamma ($\gamma : 30 - 47_{HZ}$) frequency bands using the Welch method (Kroupi et al. 2014). The Welch parameters were set as follows: frequency resolution ranged from 4 to 46 Hz, with an FFT length of 256 points and an overlap of 128 points between successive windows, to ensure sufficient spectral resolution and data continuity. These computations were performed using the default algorithms provided in the SciPy toolbox (Virtanen et al. 2020).These features provide valuable insights into brain mechanisms, and their significance has been widely investigated in the literature, establishing associations with distinct cognitive and emotional states (Ismail and Karwowski 2020).

Furthermore, in order to achieve more stable measurement and analysis, we aggregated channels based on brain regions rather than employing individual channels for the analysis of band power. Specifically, the Frontal region is constituted by averaging over PSD from AF3, AF4, F3, and F4 channels; the Right Frontal region via AF3, F3, F7, FC5, and F7 channels; the Left Frontal (LF) region via AF4, F4, FC6, and F8 channels; the Occipital region via O1 and O2 channels; and the Posterior region via O1, O2, P7, and P8 channels. Figure 2 shows the channel-to-region associations (Zanto et al. 2011).

Additionally, supplementary indices that map to the participant's cognitive state are incorporated into the feature set (Ismail and Karwowski 2020; Simon et al. 2011). Alpha spindles, $\frac{\beta}{\alpha}$ (Index 1), $\frac{\alpha+\theta}{\beta}$ (Index 2), and $\frac{\theta}{\alpha}$ (Index 3) are linked to drowsiness and can serve as basic indicators of fatigue (Chen et al. 2013; Raufi and Longo 2022; Campagne et al. 2004). Index 2 has also been associated with waning attention during prolonged tasks. Furthermore, we have incorporated $\frac{\theta}{\beta}$ (Index 4), and $\frac{\theta+\alpha}{\beta+\alpha}$ (Index 5), which are related to visual fatigue (Chen et al. 2013; Jap et al. 2009). The TLI is also added to enrich the scope of the feature set further. The TLI is calculated by the ratio of mean frontal mid-line theta energy to mean parietal alpha energy, serving as an indication for task load assessment (Berka et al. 2007; Kamzanova et al. 2014). The EEG EI, derived by calculating the ratio of beta power to the combined sum of alpha and theta powers $\frac{\beta}{\alpha+\theta}$, is introduced to evaluate task engagement and attentiveness (Prinzel et al. 2000). Finally, the feature set is augmented with theta, alpha, and beta bursts to comprehensively characterize EEG dynamics. This diversity aids in a thorough analysis of the EEG signals captured during cybersickness-inducing tasks. Table 1 provides an overview of these supplementary indices incorporated into the feature set.

Table 1 EEG indices related to drowsiness, vigilance, visual fatigue, engagement, and mental workload

Index	Formula	Interpretation
Index 1 Chen et al. (2013)	$\frac{\beta}{\alpha}$	Decrease indicates fatigue
Index 2 Chen et al. (2013) Campagne et al. (2004)	$\frac{\alpha+\theta}{\beta}$	Increase indicates fatigue and confirms attention decrease during prolonged tasks
Index 3 Raufi and Longo (2022)	$\frac{\theta}{\alpha}$	Decrease indicates fatigue and higher values suggest greater cognitive load
Index 4 Jap et al. (2009)	$\frac{\theta}{\beta}$	Increase indicates visual fatigue
Index 5 Chen et al. (2013)	$\frac{\alpha+\theta}{\alpha+\beta}$	Increase indicates visual fatigue
EI Prinzel et al. (2000)	$\frac{\beta}{\alpha+\theta}$	Measures mental engagement. Higher values indicate greater cognitive engagement
TLI Berka et al. (2007)	$\frac{F\Theta}{P\alpha}$	Quantifies mental workload. Higher values suggest greater cognitive load

3.3.2 Blood volume pulse (BVP)

The BVP signals demonstrate satisfactory SNR; however, they remain susceptible to artifacts, particularly those arising from physical movements. Hence, a band-pass filter

spanning a frequency range from 0.5 to 8 Hertz is applied to effectively remove high-frequency components. Employing the Heart Rate Analysis toolkit (Van Gent et al. 2018), further steps involve the extraction of critical cardiovascular parameters from the filtered BVP signal, including HR, Inter-Beat Interval (IBI), the Standard Deviation (Std) between adjacent beats (SDNN), Standard Deviation of Successive Differences (SDSD), Root Mean Square of Successive Differences (RMSSD), pNN20, pNN50, Median Absolute Deviation (MAD) of HR, outcomes derived from PoincarÃ© analysis (SD1, SD2, S, SD1/SD2), and estimated BR. A description of the extracted features has been provided in Table 2. This set of features enables a comprehensive analysis of the BVP signals, facilitating deep insights into cardiovascular dynamics, particularly during tasks that induce Cybersickness (Van Gent et al. 2018).

3.3.3 Electrodermal activity (EDA) and skin temperature (TMP)

The EDA signal is initially processed by employing a bandpass filter spanning a frequency range of 0.05 Hz to 4 Hz. This filtering operation is done to specifically capture the Phasic activity. Phasic activity constitutes the transient and dynamic variations in skin conductance that are reflected in the EDA signal (Boucsein 2012). Previous studies have successfully employed Empatica E4 to capture this component (Ronca et al. 2023; Pijeira-Díaz et al. 2018; Campanella et al. 2023; Collins et al. 2019). Finally, using the physio toolbox, the mean and variance of the phasic activity are calculated and added to the feature set for further analysis (Bizzego et al. 2019). Additionally, since local skin temperature is considered a reflector of ANS responses to stimuli, to further enrich the feature set, the mean and variance of the recorded skin temperature during the experimental session are incorporated (Li et al. 2022; Herborn et al. 2015). To ensure signal fidelity, a moving average technique is applied for signal smoothing prior to the inclusion of temperature-related features.

Table 2 Features derived from BVP signal and their explanations

BVP feature	Feature description
HR	A straightforward measurement of heartbeats per minute.=
BVP PSD	PSD of BVP signal between 1 and 8 Hz
MAD	A measure of heart variability
IBI	The time interval between successive heartbeats, useful for assessing HRV.
SDNN	Quantifies overall variability in heart rate, reflecting ANS function.
SDSD RMSSD	SDSD and RMSSD analyze the variability of successive heartbeat intervals, with RMSSD particularly sensitive to parasympathetic activity
pNN20 pNN50	Percentages of differences between adjacent R-R intervals,greater than 20 ms and 50 ms, respectively, indicators of rapid changes in heart rate.
SD1	Std of points perpendicular to the line of identity, indicating short-term HR variability.
SD2	Standard deviation along line of identity, reflecting long-term variability.
S	Quantifies scatter of points within plot, indicating overall heart rate variability.
SD1/SD2 ratio	Provides insights into balance between short-term and long-term heart rate variabilities, reflecting relative contributions of sympathetic and parasympathetic inputs to heart rate dynamics.
BR	Estimated breathing rate extracted from BVP signal

4 Data analysis

In this section, we share the outcomes of our experiments. Initially, we provide a detailed overview of the participants (Subsection 4.1). We analyze the SSQ responses in Subsection 4.2 to determine the total SSQ score.

In Subsection 4.3, we present the outcomes of statistical analyses conducted on the processed physiological data, identifying parameters that are significantly distinct across different stages of the experiment and between the primary groups.

4.1 Participants

A total of 24 participants, comprising 4 women and 20 men with ages ranging from 15 to 52, participated in the simulation experiment. The majority of the participants fell within the 15 to 25 age bracket, and 60% of the participants had prior experience with VR. All participants were healthy and had not experienced any brain-related injuries. They had normal or corrected-to-normal vision. This study was conducted as part of the WaveVR project, with ethical approval from Ghent University. Moreover, the participants were asked for explicit consent at the beginning of the sessions following the GDPR rules.

It is crucial to mention that, to enhance the SNR and ensure physiological signal integrity, a visual inspection is conducted on the recorded signals of each participant. Consequently, physiological data acquired from two participants was omitted from the analysis due to noise contamination attributed to pronounced movements, resulting in 22 participants in the final dataset.

The experimental setup involved a VR environment wherein participants interacted with a roller coaster simulation. Prior to starting the session, we ensured that every participant received detailed information about the procedural steps. Additionally, we requested that they inform us if they had any pre-existing issues related to their cardiovascular

or visual health. Based on the responses received, we found that no relevant issues were reported by the participants.

4.2 Analysis of simulator sickness questionnaire (SSQ)

Figure 6 provides a visual representation of the scores derived from the SSQ, which were calculated based on the formula provided by Kennedy et al. (1993) . These scores are grouped into three scales: nausea, which gauges feelings of nausea; oculomotor, encompassing factors related to vision; and disorientation, which pertains to feelings of being directionally unsettled. In the assessment of the 16 symptoms, general discomfort displays the highest average score of 1.4 on a scale of 0 (absence) to 3 (high severity). Other symptoms, such as nausea, dizziness (eyes closed), vertigo, and stomach awareness, also show similar patterns with average scores exceeding 1.

It is noteworthy that significant score variations are observed in relation to the disorientation scale. Moreover, Nausea is reported more frequently than disorientation. Analyzing the overall score distribution highlights the divergence in the extent of Cybersickness experienced across individuals. This also highlights the complexity of detecting Cybersickness through subjective reports, as it is highly dependent on the individual's perception.

With the aim to cluster behavioral patterns, the total SSQ scores are individually plotted for each test subject, as depicted in Fig. 7. The SSQ scores, which can range from 0 to 200, showed significant variability among participants. The reported minimum and maximum scores were 7.48 and 142.12, respectively, indicating a wide range of severity in Cybersickness symptoms. These results demonstrate the considerable heterogeneity in the reported symptom severity. The mean SSQ score for this experiment equals 63.26. The visualization in Fig. 7 provides a clear distinction between subjects with elevated SSQ scores and those with lower scores.

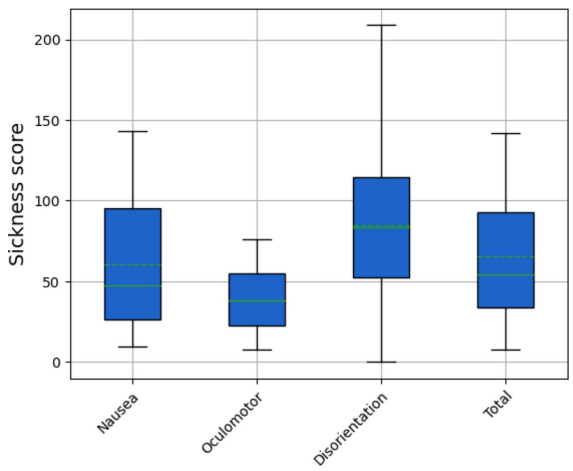


Fig. 6 Distribution of SSQ scores across three criteria: Nausea, Oculomotor, and Disorientation

To further analyze the data, the total SSQ scores are partitioned into two distinct clusters using the K-means algorithm (Pedregosa et al. 2011). The first cluster, with a mean score of 34.81, is categorized as the low Cybersickness symptom group, while the second cluster, with a mean score of 96.9, is identified as the high Cybersickness symptom group. Additionally, one subject with a total SSQ score of 71.06 is excluded to widen the margin between the two groups. Consequently, 8 subjects are allocated to the high Cybersickness symptoms group, and 13 subjects are assigned to the low Cybersickness symptoms group. Henceforward, the analysis will be conducted in accordance with these two defined groups.

4.3 Statistical analysis of physiological data versus subjective reports

In this study, the primary aim is to investigate the correlation between physiological metrics and Cybersickness level, as quantified by the SSQ. Given the potential for the experimental paradigm, specifically the roller coaster simulation, to induce physiological alterations, it is imperative to first identify and control for these variables.

We conducted three sets of statistical tests to address the influence of the VR roller coaster simulation on physiological parameters across participants and their relationship with Cybersickness. Initially, we performed a repeated measures statistical test on the entire dataset, regardless of the reported SSQ scores, to examine within-subject effects over time across the different stages of the experiment. To better understand the specific impact of Cybersickness on physiological responses, we conducted two additional sets of repeated measures tests, focusing on participants grouped by their reported levels of Cybersickness. This categorization allows us to isolate and analyze the physiological effects attributed uniquely to different Cybersickness intensities. By comparing groups with elevated Cybersickness scores to those with lower scores, we sought to identify distinct physiological patterns and determine how Cybersickness contributes to overall variance in the dataset.

For these comparisons, we employed the Friedman test, a non-parametric statistical technique used to analyze differences when the same subjects are exposed to different stimuli (Sheldon et al. 1996). In our case, these stimuli correspond to varying levels of motion complexity during the experiment. The Friedman test helps determine whether there are statistically significant differences in measures between different experiment phases. In each analysis, the physiological feature under scrutiny served as the dependent

Fig. 7 Scatter plot of individual SSQ scores, indicating variability and mean score across participants

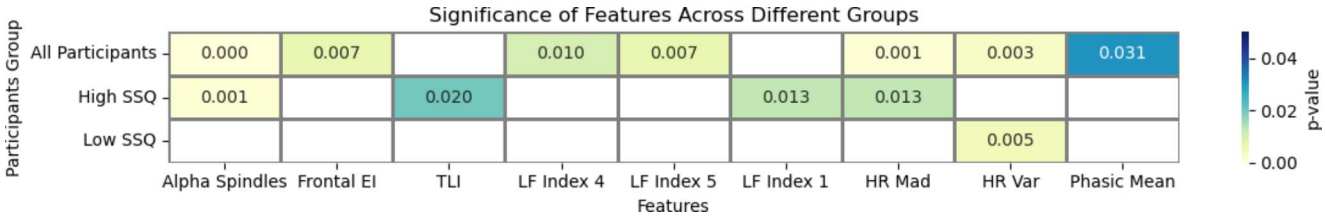


Fig. 8 Distribution of parameters exhibiting significant variations among participants: Values inside each cell represent the p-value obtained from statistical tests, with white cells indicating a lack of statistical significance

variable, while the temporal phase of the experiment was the independent variable. When the Friedman test indicated significant changes, we performed a Nemenyi post-hoc test to identify which pairs were significantly different from each other (Terpilowski 2019). These comparisons included pre-stimulus vs. stimulus, pre-stimulus vs. post-stimulus, and stimulus vs. post-stimulus.

The first set of tests examined the impact of the VR simulation on all participants, identifying significant changes in several EEG and physiological features. The most notable changes included the alpha spindle, Index 4 and 5 in the LF region, the EI in the same region, and MAD and HR variance. As depicted in Fig. 9a, alpha spindle activity decreased during the VR simulation. Post-hoc analysis pinpointed these changes as specifically significant between the post-stimulus and stimulus phases. This reduction likely signifies increased vigilance, suggesting that the VR simulation induces user alertness. This pattern was consistent

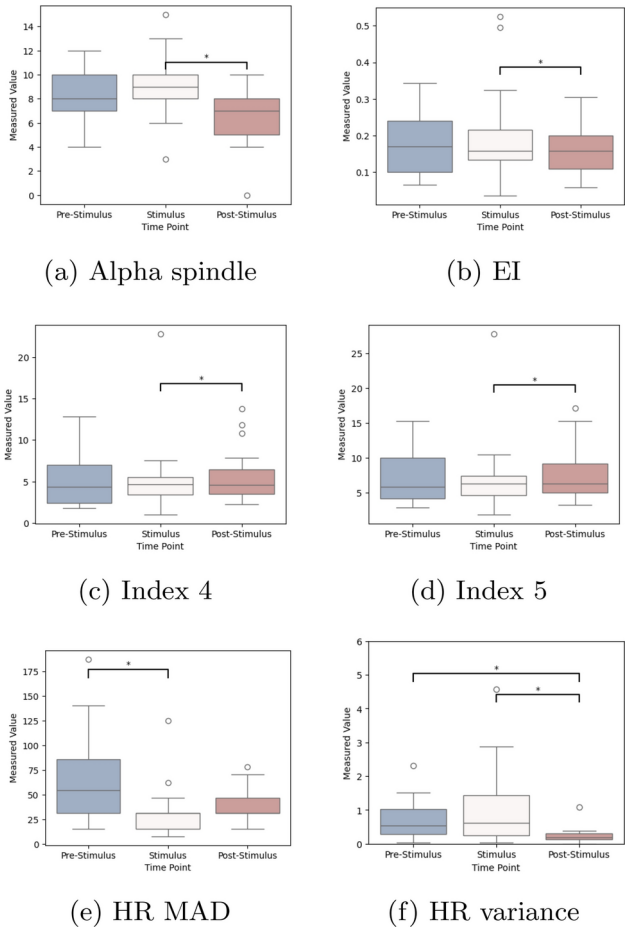


Fig. 9 six significant features across three experimental phases. The y-axis represents the magnitude of the physiological response. The x-axis represents the experiment phase

across both groups of participants, aligning with fluctuations observed in the frontal EI. Notably, the EI remained relatively stable in the pre- and post-stimulus phases but showed an increase during the stimulus phase, as illustrated in Fig. 9b. Similar to the alpha spindles, significant changes in the EI were noted between the post-stimulus and stimulus phases.

The analysis of EEG-derived indices related to visual fatigue, specifically Index 4 and 5, revealed significant increases during and after the stimulation phase, particularly among participants with high Cybersickness scores. Index 4 did not show significant changes between the pre-stimulus and stimulus phases; however, a notable increase was observed during the post-stimulus phase compared to the stimulus phase. This trend suggests that visual fatigue accumulates over the duration of the VR experience and becomes more pronounced post-simulation. Moreover, the elevation in this index was significantly more marked in the group reporting high levels of Cybersickness, indicating a potential correlation between the severity of Cybersickness symptoms and the extent of visual fatigue experienced. These changes are visually represented in Fig. 9c, highlighting the differential impact across the experimental phases. Similarly, Index 5 exhibited slight increases during the stimulus phase and more substantial rises in the post-stimulus phase, as depicted in Fig. 9d. Analysis of data derived from BVP and EDA measurements revealed significant physiological responses to the VR simulation across participants. Key features showing noteworthy changes included HR MAD and HR variance, depicted in Fig. 9. The HR MAD significantly decreased during the stimulus phase compared to the pre-stimulus phase. This indicates a physiological response possibly related to initial adaptation or a parasympathetic response to high-complexity VR motion.

The second set of analyses identified several features that significantly changed among participants who reported higher levels of Cybersickness. These features include Index 1 in the LF regions, TLI, MAD, SD2, and HR mean. A decrease in Index 1, typically associated with fatigue, was observed. Post-hoc analysis indicated significant differences in this index between the pre-stimulus and stimulus phases, as well as between the post-stimulus and stimulus phases in the LF regions. Notably, this decrease was prevalent among almost all participants with higher levels of Cybersickness, suggesting a strong correlation between Cybersickness severity and fatigue levels. Conversely, participants with lower levels of Cybersickness did not exhibit a consistent pattern in this index, highlighting differential physiological impacts based on individual susceptibility to Cybersickness. This finding initially appears to contradict the increased alertness suggested by alpha spindle activity observed across all participants. Upon closer examination, the decrease in the Index 1 ratio, indicating diminished beta activity relative to alpha, can be interpreted within the specific context of Cybersickness. Despite a decrease in alpha activity, there was a reduction in beta activity, likely due to the increased cognitive and physical strain associated with managing Cybersickness symptoms.

Furthermore, a significant decrease in the TLI was observed from pre-stimulus to post-stimulus in participants who reported high levels of Cybersickness. This trend was not observed in participants with lower levels of Cybersickness, pointing to a divergence in cognitive burden between the two groups. This decrease in TLI among the high Cybersickness group suggests that these participants experienced a reduction in cognitive engagement over time, potentially as a response to the discomfort or sensory overload caused by the VR environment. This finding aligns with the observation for Index 1, which indicated increased cognitive strain and fatigue in participants experiencing higher cybersickness.

The third set of analyses revealed that HR variance was the sole feature exhibiting significant changes among participants who reported lower levels of Cybersickness. Specifically, HR variance increased during the stimulus phase and subsequently decreased in the post-stimulus phase. Notably, significant differences were observed between pre-stimulus and post-stimulus measurements. The increase in HR variance during the stimulus phase suggests a heightened autonomic response to the VR environment, indicative of increased physiological arousal. This response is expected as participants engage with VR stimuli, reflecting the normal reaction to engaging scenarios.

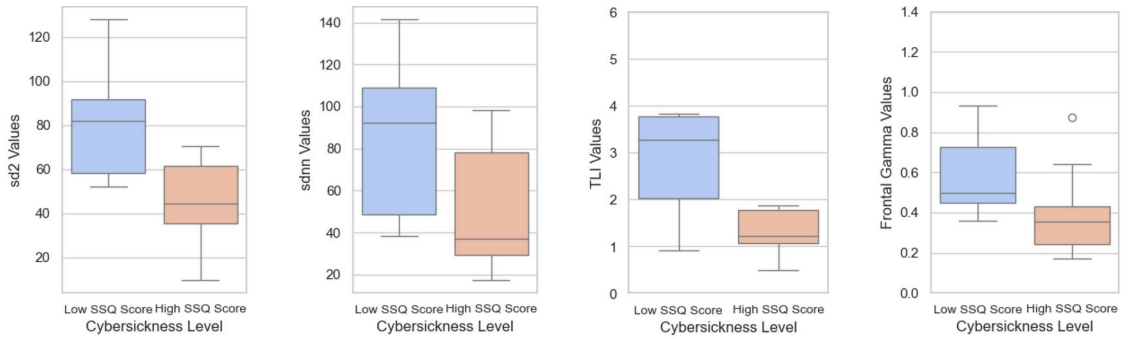
Furthermore, we performed a set of Mann–Whitney tests with Bonferroni corrections applied for each simulation phase to identify significant differences between the high and low SSQ groups across the experimental phases. This test is particularly useful for comparing distributions between two independent groups, especially when distributions are non-normal. No significant features were identified between the two groups after correction. However, during the stimulus phase, several features extracted from the BVP signal showed p-values less than 0.05 prior to Bonferroni correction, including SDNN, SD2, and S. These features exhibited higher values in the group with lower Cybersickness symptoms. As illustrated in Fig. 10, the group with higher levels of Cybersickness symptoms demonstrated lower levels of SDNN, SD2, and S, indicating reduced long-term HRV. Conversely, the group with lower levels of Cybersickness symptoms showed elevated levels of these features. This divergence may be attributed to the heightened vigilance of the lower-symptom group during the experiment, enabling them to respond more adaptively to the roller coaster simulation and thereby manifest higher HRV. In the post-stimulus phase, several features showed p-values less than 0.05 prior to Bonferroni correction, including TLI, LF Index 1, and frontal gamma activity. Frontal gamma activity, associated with higher cognitive processes and attention, varied significantly between the two groups in the post-stimulus phase. This finding is consistent with previous analyses, indicating that VR-induced stress and cognitive load may lead to fluctuations in brain activity associated with active cognition and attention.

Conducting statistical tests enables us to elucidate the relationship between individual physiological features and reported levels of Cybersickness symptoms. However, it is important to acknowledge that the effects of changes in the human nervous system, particularly in the sympathetic and parasympathetic branches, may not manifest in a linear fashion. The Friedman test identifies overall differences in ranks across multiple related groups, which may not capture complex interactions between features and Cybersickness levels. Furthermore, a unidimensional perspective on the data would thus constrain our ability to uncover more intricate interrelationships between features, thereby limiting the efficacy of Cybersickness detection or prediction. To address this limitation, the study aims to explore more complex relationships and enhance the accuracy of Cybersickness symptom detection by employing predictive models based on physiological recordings.

5 Cybersickness modeling

The purpose of this section is to to construct a predictive model for Cybersickness utilizing a classifier predicated on a feature set analogous to those examined in preceding sections. To address Cybersickness challenges, physiological signals are utilized for the modeling of Cybersickness levels.

Figure 11 provides an overview of a system designed for Cybersickness detection and feedback provision. The model is trained on both physiological and subjective data collected during the training phase, which is discussed in subsections 5.1 and 5.2. The model serves a pivotal role in



(a) SD2 during stimulus phase (b) SDNN during stimulus phase (c) TLI index during post-stimulus phase (d) γ activity during post-stimulus phase

Fig. 10 Box plot of four of the significant features differentiating low and high Cybersickness symptom groups.The y-axis represents the magnitude of physiological response relative to Low and high SSQ scores

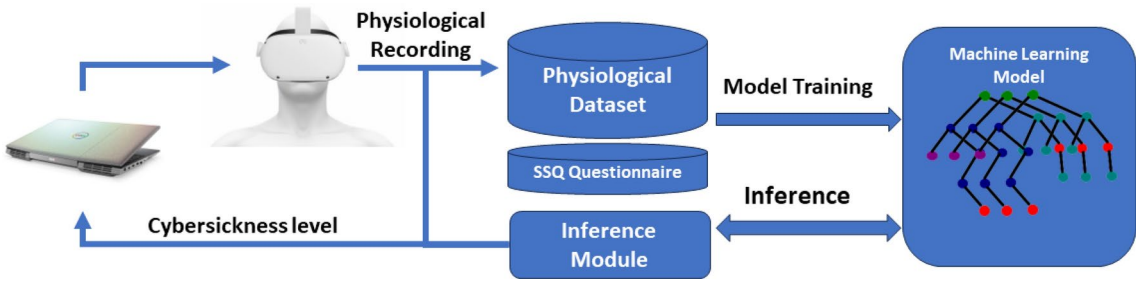


Fig. 11 Schematic overview of the modeling workflow for Cybersickness training and inference

the system by providing feedback, thereby enabling system to trigger appropriate interventions to mitigate Cybersickness. During the evaluation and inference stages, new physiological data are input into the model to infer the level of Cybersickness. An estimation of the performance of such a model is discussed in subsection 5.3. This information is then relayed back to the rendering system to initiate an appropriate response.

Subsequent sections detail the data preprocessing methods, describe the utilization of five modeling techniques, and present their respective evaluations.

5.1 Preprocessing

We incorporate data from all four signal modalities for model construction. The preprocessing steps for BVP, EDA, and TMP are done according to the approach delineated in section 3.3. Furthermore, to standardize the scale of features across different modalities and enhance the robustness of our analysis, we employed a robust scaler to normalize them (Pedregosa et al. 2011). This scaler is particularly effective as it is less sensitive to outliers in the data, ensuring that extreme values do not disproportionately influence the model's performance.

5.2 Modeling

We employ five classifiers– Support Vector Machine (SVM), Decision Tree (DT), Linear Discriminant Analysis (LDA), Logistic Regression (LR), and XGBoost–to assess their efficacy in predicting or detecting Cybersickness in a binary setting. This approach simplifies the modeling process and focuses on the discriminative capacity of the physiological signals(Parui et al. 2019; Khoirunnisaa et al. 2018). Each classifier is applied in three distinct experimental settings. Initially, we evaluate performance using features solely derived from brain-related modality EEG data as an input. Subsequently, we assess the classifiers while only feeding features extracted from body-related modalities, including BVP, EDA, and TMP. Lastly, we combine features from both brain and body modalities for a comprehensive classification. Training classifiers on individual physiological modalities reduces model complexity and dimensionality. Conversely, integrating multiple modalities offers a comprehensive perspective on physiological states, thereby enabling the model to discover inter-modal correlations. Additionally, while data from brain modalities are more susceptible to motion artifacts than body modalities, these limitations can be reduced by incorporating additional information from body modalities.

To ensure robust generalizability assessments, we partitioned the dataset into five subsets and conducted the experiment iteratively, employing a 5-fold cross-validation approach with 10 repetitions. For SVM, the models were instantiated with the C parameter set to 1, utilizing a Radial Basis Function (RBF) kernel. DT models were configured using the Gini criterion for node splitting. LR models had a regularization strength set to 1 with an L2 penalty term. XGBoost models were deployed with a Gbtree booster and a gamma parameter fixed at 1. These parameter settings were chosen based on their common use as default values in the literature, which provide a balance between performance and computational efficiency for preliminary analyses (Chen and Guestrin 2016; Pedregosa et al. 2011). Given the scope of our dataset and the exploratory nature of this study, these settings were deemed appropriate to minimize the risk of overfitting while maintaining computational feasibility.

5.3 Results

Table 3 provides a comprehensive evaluation of the performance of multiple machine learning classifiers, in a binary classification setting, distinguishing between participants experiencing high and low levels of Cybersickness. Specifically, the performance of SVM, DT, LDA, LR, and XGBoost classifiers are assessed across three distinct data modalities: brain, body, and a combination of both.

The evaluation metrics employed in this study comprise accuracy, F1 Score, and coefficient of variation (CV). Accuracy measures the proportion of correct predictions out of all predictions made, providing a basic indicator of overall

Table 3 Binary classification results from SVM,DT,LDA, LR, and XGBoost

Modality	Brain			Body			Both		
Classifier	Accuracy	F1	CV	Accuracy	F1	CV	Accuracy	F1	CV
SVM	74.61	73.59	14.62	67.58	65.27	13.53	73.33	70.06	15.45
DT	72.23	70.77	15.48	75.92	74.86	14.25	75.77	75.14	12.33
LDA	68.07	68.01	15.73	64.72	64.04	**13.28**	68.88	68.94	27.74
LR	**77.69**	**77.58**	18	65.93	64.33	23	75.55	75.38	17
XGBoost	76.41	75.22	**14.03**	**79.45**	**79.29**	14.68	**86.66**	**86.35**	**9.5**

The best values for each modality are shown in bold.

performance. The F1 Score, the harmonic mean of precision and recall, effectively balances the trade-off between false positives and false negatives, making it ideal for imbalanced datasets. Precision is the proportion of correct positive predictions among all positive predictions made, while recall tracks actual positives identified correctly. Thus, the F1 Score merges these into a single metric for a comprehensive evaluation of classifier performance. The CV assesses the stability and reliability of the models by measuring the relative variability of their performance across different data subsets. Specifically, a lower CV indicates less variation in model performance, suggesting more stable results across subsets and implying better generalization within the dataset. These metrics are selected to comprehensively assess each classifier's predictive performance and generalizability. False positives are undesirable as they can disrupt user immersion by triggering unnecessary interventions, such as displaying a fixed point on the screen when the user is not actually experiencing Cybersickness. On the other hand, false negatives are equally detrimental, as they compromise the Quality of Experience (QoE) by failing to recognize and address the user's discomfort.

In the Brain modality, LR emerges as the most effective classifier, registering an accuracy of 77.69%, an F1 score of 77.58%, and a CV score of 18. This indicates that although LR has an average good accuracy, it does not generalize well across the dataset because of the high CV score when classifying high and low levels of Cybersickness based solely on brain-related features. XGBoost is the next most effective, with an accuracy of 74.41% and an F1 score of 74.22%; it still exhibits a high CV score of 14.03, indicating instability across different data subsets.

In the Body modality, XGBoost surpasses all competing classifiers, registering an accuracy of 79.45% and an F1 score of 79.29%. These metrics indicate that XGBoost serves as the most reliable classifier for differentiating between high and low Cybersickness levels when solely relying on body-related physiological features. DT also demonstrates commendable performance in this domain, achieving an accuracy of 75.92% and an F1 score of 74.86%. Given that both XGBoost and DT are tree-based models, their efficacy may suggest the inherent advantages of tree-based algorithms in classifying Cybersickness based exclusively on body-related modalities.

However, the most striking results are observed in both modalities, which incorporate features from both Brain and Body data. Here, XGBoost significantly outshines all other classifiers, achieving an exceptional accuracy of 86.66% and an F1 score of 86.35%. Moreover, it demonstrates the lowest CV of 9.5, indicating the highest level of performance stability across different data subsets. The superior performance of XGBoost in both modalities suggests that a multimodal approach, incorporating both brain and body-related features, offers the most effective and reliable method for classifying levels of Cybersickness. This could have significant implications for the development of more accurate and robust Cybersickness detection systems.

6 Explainable AI (XAI) for model interpretation

In this section, we outline our utilization of the SHapley Additive exPlanations (SHAP) framework, a state-of-the-art method for model interpretability, in order to enhance transparency and facilitate understanding of the intricate relationships between physiological markers and subjective experiences (Lundberg and Lee 2017; Lundberg et al. 2018). This interpretability not only fosters trust in the model's predictive capabilities but also aids researchers in tailoring more effective interventions and mitigation techniques to prevent Cybersickness.

Specifically, focus is placed on XGBoost models as they outperform other classification techniques in terms of accuracy. Approaches with suboptimal predictive performance are excluded from this interpretive analysis to ensure the credibility of the insights derived. Thus, we employed 50 XGBoost models, trained during the 10-5-fold cross-validation, for subsequent analysis. Furthermore, due to the time consumption and computational intensity of conducting SHapley Additive exPlanations (SHAP) analysis on each of the XGBoost models individually, an aggregate analysis is opted for. More specifically, a majority rule ensemble of all classifiers were constructed and then SHAP analysis were performed on this ensemble to provide a more balanced

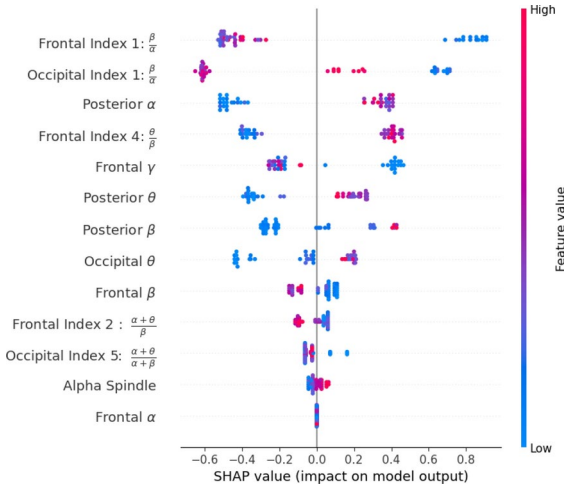


Fig. 12 Summary plot of SHAP values for features extracted from EEG signals

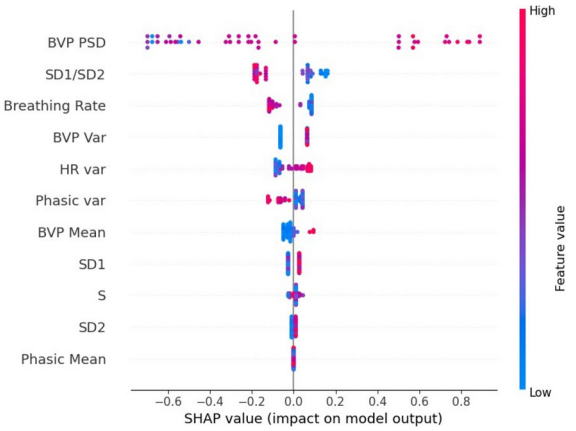


Fig. 13 Summary plot of SHAP values for features extracted from BVP, EDA, and TMP signals

representation of feature importance, thereby mitigating scaling issues that could arise from individual model variations. Using this method, a deeper understanding of feature contributions and models' behaviors is gained, resulting in more accurate and reliable findings.

Figures 12 and 13 depict feature contributions using a SHAP summary plot. A SHAP summary plot offers a comprehensive view of feature importance across the test set. Each data point in the plot signifies a SHAP value for a specific feature in a given instance. The x-axis quantifies the magnitude of the SHAP value, indicating the feature's influence on models predictions, while the y-axis enumerates the features, typically sorted by their aggregate importance. The color gradient serves as an indicator of intensity values: red denotes higher values, whereas blue signifies lower values for each data point.

As illustrated in Fig. 12, the features are systematically ranked based on their relative importance. Upon initial examination of the summary plot, it becomes evident that a complex interplay exists among the features in determining the model output. Notably, the values of most features do not serve as absolute indicators of either high or low Cybersickness symptoms. However, certain features manifest a significant correlation with the levels of Cybersickness.

Prominently, Index 1, extracted from the frontal region and identified as an indicator of both vigilance and drowsiness, emerges as the most salient feature (Chen et al. 2013). Elevated values of this index are inversely correlated with an increased likelihood of subjects manifesting Cybersickness symptoms. In essence, subjects in a state of heightened drowsiness are more susceptible to experiencing Cybersickness. This observation is corroborated by the second most salient feature, Index 1, extracted from the occipital region, thereby reinforcing the aforementioned interpretation consistent with our analysis in section 4.

Another factor that further supports the previous observation is the third contributing factor found by the model, which is alpha activity in the posterior region. It is widely recognized that an escalation in the PSD of alpha frequency in the posterior region is an indicative sign of drowsiness (Cantero et al. 2002). As the SHAP value of this feature indicates, individuals susceptible to Cybersickness demonstrate elevated levels of alpha activity in the posterior region.

Following this, Index 4, also extracted from the frontal region, emerges as a salient feature. This index is directly associated with visual fatigue, and the model tends to assign subjects to the high Cybersickness group if they exhibit elevated values in Index 4, which is in line with our analysis in section 4 (Jap et al. 2009; Ismail and Karwowski 2020). Furthermore, the model establishes a non-uniform reverse correlation between the level of gamma activity and reported Cybersickness symptoms in line with our observation in section 4 (see Fig. 10d). This non-uniformity arises from the multivariate effects of features, indicating that the model's decisions are not influenced by a single feature. In some instances, lower values of gamma activity positively affect the model's output to predict elevated cybersickness, but this does not hold for all instances. This finding warrants further investigation of multivariate analysis of the relationship between gamma activity and Cybersickness symptoms across a broader spectrum.

Moreover, although not universally applicable to each instance, the model tends to attribute higher Cybersickness levels to individuals exhibiting elevated theta activity in the occipital and posterior regions, a phenomenon positively correlated with increased fatigue (Ismail and Karwowski 2020; Lin et al. 2014; Tanaka et al. 2012). Subsequently, the next feature of significance is beta activity extracted from the posterior region. Lower Beta activity in this region predominantly influences the model's decision to categorize

subjects into the low SSQ score group. Elevated beta activity in the posterior region has been previously associated with visual attention.

Additionally, another feature of note is beta activity extracted from the frontal region. The model is inclined to categorize subjects into the high Cybersickness group if they display reduced beta PSD. Lower beta PSD has been previously linked to mental fatigue, which is compatible with our finding regarding LF Index 1 (Zhang and Yu 2010; Lin et al. 2014; Chen et al. 2013). Furthermore, the emergence of these features indicates a potential link between the Sopite syndrome, a neurological disorder characterized by fatigue and drowsiness during prolonged motion exposure, and Cybersickness. This association highlights the need for additional research to explore their common physiological and psychological underpinnings Lawson and Mead (1998))

Furthermore, it is observed that the model assigns a negative correlation to Index 2 in the frontal region. This index gradually decreases over prolonged task durations (Ismail and Karwowski 2020). Conversely, this index is inversely related to the EI, implying that lower values of Index 2 and higher values of the EI in this region tend to predict high Cybersickness symptoms. It is conceivable that individuals without Cybersickness symptoms show a normal increase in this proportion during prolonged task engagement, which presumably doesn't occur in subjects with high Cybersickness symptoms. This finding necessitates further exploration regarding task engagement and attention study. Additionally, Index 5, known to directly correlate with visual fatigue, also appears as an important feature, though instances of distinct differences between high and low values are noted. A clear relationship between elevated values of Index 5 and high Cybersickness symptoms is evident.

Turning to the features associated with body modalities, Fig. 13 systematically ranks these features based on their relative importance. A similar complex interplay among features is observed in determining the model's output, akin to what was noted for brain-related modalities. The most salient feature, according to the model's perspective, is the PSD of the BVP signal. This characteristic may suggest a correlation between mental stress levels and cybersickness, as fluctuations in the BVP have previously been linked to mental stress (Cipresso et al. 2019; Zhai et al. 2005).

The relationship between the $\frac{SD1}{SD2}$ values and the model's outcome reveals a negative correlation, suggesting that a reduction in this metric tends to result in the classification of high Cybersickness symptoms. SD1 predominantly captures short-term variability, while SD2 assesses long-term variability in the data. It's worth highlighting that this feature is an indicator of a change of balance in autonomic nervous system activity (Shaffer and Ginsberg 2017). A reduction in the ratio of SD1 to SD2 may indicate a shift towards sympathetic dominance, which is indicative of a stress response. This interpretation aligns with the inclination of the model to classify individuals with high Phasic activity variance as those who experienced a mild level of Cybersickness. However, in our specific case, where each recording phase is limited to less than a minute, it is advisable to conduct further investigations to discover the full implications of these findings in brief recordings.

Interestingly, the model output is inversely influenced by BR; subjects experiencing elevated symptoms of Cybersickness tend to have a reduced BR. It's important to note that this negative correlation with BR could be contextually influenced by the experimental environment. For example, subjects with more severe symptoms might experience altered BR due to the anxiety induced by the roller coaster's fluctuations.

In summary, the high-accuracy model allocates substantial weights to both brain and body modalities, incorporating indices derived from EEG as well as raw brain power rhythms. Notably, these indices–particularly those associated with drowsiness, fatigue, and alpha spindle occurrence–emerge as robust predictors of Cybersickness. The integration of these indices with features derived from body modalities holds promise for the advancement of real-time Cybersickness detection systems, complementing the utility of raw power rhythms.

7 Conclusion

In this paper, we explored the interrelation between physiological signals and self-reported scores to detect and quantify Cybersickness within VR environments. Through the deployment of a VR roller coaster simulation, our objective was to establish robust connections between physiological signals and self-reported SSQ ratings. Our findings not only revealed changes in physiological signals attributable to the simulation but also identified significant alterations in the group with elevated SSQ scores.

Building on these insights, we constructed a predictive model based on physiological markers, notably achieving an accuracy rate of 86.66%. These preliminary outcomes indicate the promising potential of physiological signals to supplant traditional questionnaires in assessing Cybersickness. This shift holds promise for developing real-time, non-disruptive assessment methods, opening the door for designing and implementing more immersive and adaptive systems.

Despite the robust performance of XGBoost in classifying Cybersickness levels, the obscure nature of such machine learning models limits understanding of the underlying physiological processes. Our application of XAI

techniques was an attempt to explain these models by highlighting influential features, but ongoing efforts are needed to refine these explanations and further align them with established physiological theories.

Future research could broaden the scope to include a more diverse array of physiological data, such as eye tracking, pupillometry, and muscle activity. Additionally, implementing detailed pre- and post-stimulus questionnaires could yield models capable of discerning not only the presence of Cybersickness but also its severity and progression over time.

Furthermore, addressing the demands of real-world applications necessitates expanding participant diversity and diversifying the simulations employed. Although a static setting was essential in this study to mitigate motion-induced artifacts, future research should consider dynamic, ambulatory settings that allow participants to move freely. This approach would enhance the ecological validity of the findings and enable a more comprehensive analysis of environmental factors influencing Cybersickness. The increased dataset size from such studies could facilitate the development of sophisticated, potentially real-time predictive models that adapt more dynamically to individual user responses.

In summary, this paper serves as a step toward a more comprehensive, multimodal approach for the quantitative assessment of Cybersickness. It not only enhances the objectivity of Cybersickness assessment but also lays the groundwork for future research aimed at improving user experience in immersive technologies.

Acknowledgements This research is partly funded by the FWO WaveVR project (Grant Number: G034322N). Sam Van Damme is also funded by FWO (Grant Number 1SB1822N).

Competing interests The authors declare no competing interests.

References

Ahn MH, Park JH, Jeon H et al (2020) Temporal dynamics of visually induced motion perception and neural evidence of alterations in the motion perception process in an immersive virtual reality environment. Front Neurosci 14:600839

Asl NS, Baghdadi G, Ebrahimian S et al (2022) Toward applicable EEG-based drowsiness detection systems: a review. Front Biomed Technol. https://doi.org/10.18502/fbt.v9i4.10426

Awais M, Badruddin N, Drieberg M (2017) A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability. Sensors 17(9):1991

Berka C, Levendowski DJ, Lumicao MN et al (2007) EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. Aviat Space Environ Med 78(5):B231–B244

Bizzego A, Battisti A, Gabrieli G et al (2019) pyphysio: a physiological signal processing library for data science approaches in physiology. SoftwareX 10:100287

Borghini G, Astolfi L, Vecchiato G et al (2014) Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci Biobehav Rev 44:58–75

Boucsein W (2012) Electrodermal activity. Springer Science & Business Media, Berlin

Campagne A, Pebayle T, Muzet A (2004) Correlation between driving errors and vigilance level: influence of the driver's age. Physiol Behav 80(4):515–524

Campanella S, Altaleb A, Belli A et al (2023) A method for stress detection using empatica E4 bracelet and machine-learning techniques. Sensors. https://doi.org/10.3390/s23073565

Cantero JL, Atienza M, Salas RM (2002) Human alpha oscillations in wakefulness, drowsiness period, and rem sleep: different electroencephalographic phenomena within the alpha band. Neurophysiol Clin Clin Neurophysiol 32(1):54–71

Cevette MJ, Stepanek J, Cocco D et al (2012) Oculo-vestibular recoupling using galvanic vestibular stimulation to mitigate simulator sickness. Aviat Space Environ Med 83(6):549–555. https://doi.org/10.3357/ASEM.3239.2012

Chang E, Billinghurst M, Yoo B (2023) Brain activity during cybersickness: a scoping review. Virtual Real 27:1–25

Charbonnier S, Roy RN, Bonnet S et al (2016) EEG index for control operators' mental fatigue monitoring using interactions between brain regions. Expert Syst Appl 52:91–98

Chen C, Li K, Wu Q et al (2013) EEG-based detection and evaluation of fatigue caused by watching 3DTV. Displays 34(2):81–88

Chen T, Guestrin C (2016) XGBoost: ascalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp 785–794

Cipresso P, Colombo D, Riva G (2019) Computational psychometrics using psychophysiological measures for the assessment of acute mental stress. Sensors. https://doi.org/10.3390/s19040781

Clifton J, Palmisano S (2020) Effects of steering locomotion and teleporting on cybersickness and presence in HMD-based virtual reality. Virtual Real 24(3):453–468

Cohen MX (2014) Analyzing neural time series data: theory and practice. MIT Press, Cambridge

Collins J, Regenbrecht H, Langlotz T et al (2019) Measuring cognitive load and insight: a methodology exemplified in a virtual reality learning context. In: 2019 IEEE International symposium on mixed and augmented reality (ISMAR). IEEE, pp 351–362

Delorme A (2023) EEG is better left alone. Sci Rep 13(1):2372

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134(1):9–21

Dennison MS, Wisti AZ, D'Zmura M (2016) Use of physiological signals to predict cybersickness. Displays 44:42–52
Doran D, Schulz S, Besold TR (2017) What does explainable AI really mean? A new conceptualization of perspectives. arXiv preprint arXiv:1710.00794
Emotiv (2024) Emotive epoch x. https://www.emotiv.com/epoc-x/
Empatica (2024) Empatica e4. https://www.empatica.com/research/e4/
Fernandes AS, Feiner SK (2016) Combating VR sickness through subtle dynamic field-of-view modification. In: 2016 IEEE symposium on 3D user interfaces (3DUI), pp 201–210. https://doi.org/10.1109/3DUI.2016.7460053
Gans J, Nagaraj A (2023) The economics of augmented and virtual reality. arXiv preprint arXiv:2305.16872
Garcia-Agundez A, Reuter C, Becker H et al (2019) Development of a classifier to determine factors causing cybersickness in virtual reality environments. Games Health J 8(6):439–444
Garcia-Agundez A, Reuter C, Caserman P et al (2019) Identifying cybersickness through heart rate variability alterations. Int J Virtual Real 19(1):1–10
Garrido LE, Frías-Hiciano M, Moreno-Jiménez M et al (2022) Focusing on cybersickness: pervasiveness, latent trajectories, susceptibility, and effects on the virtual reality experience. Virtual Real 26(4):1347–1371
Groth C, Tauscher JP, Heesen N et al (2022) Omnidirectional galvanic vestibular stimulation in virtual reality. IEEE Trans Vis Comput Graph 28(5):2234–2244. https://doi.org/10.1109/TVCG.2022.3150506
Guna J, Geršak G, Humar I et al (2019) Influence of video content type on users' virtual reality sickness perception and physiological response. Futur Gener Comput Syst 91:263–276
Herborn KA, Graves JL, Jerem P et al (2015) Skin temperature reveals the intensity of acute stress. Physiol Behav 152:225–230
Hirzle T, Cordts M, Rukzio E et al (2021) A critical assessment of the use of SSQ as a measure of general discomfort in VR head-mounted displays. In: Proceedings of the 2021 CHI conference on human factors in computing systems, pp 1–14
Islam R, Ang S, Quarles J (2021) CyberSense: a closed-loop framework to detect cybersickness severity and adaptively apply reduction techniques. In: 2021 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW), pp 148–155. https://doi.org/10.1109/VRW52623.2021.00035
Ismail LE, Karwowski W (2020) Applications of EEG indices for the quantification of human cognitive performance: a systematic review and bibliometric analysis. PLoS ONE 15(12):e0242857
Jap BT, Lal S, Fischer P et al (2009) Using EEG spectral components to assess algorithms for detecting fatigue. Expert Syst Appl 36(2):2352–2359
Kamzanova AT, Kustubayeva AM, Matthews G (2014) Use of EEG workload indices for diagnostic monitoring of vigilance decrement. Hum Factors 56(6):1136–1149
Kennedy RS, Lane NE, Berbaum KS et al (1993) Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psychol 3(3):203–220
Kennedy RS, Drexler J, Kennedy RC (2010) Research in visually induced motion sickness. Appl Ergon 41(4):494–503. https://doi.org/10.1016/j.apergo.2009.11.006
Keshavarz B, Golding JF (2022) Motion sickness: current concepts and management. Curr Opin Neurol 35(1):107–112. https://doi.org/10.1097/WCO.0000000000001018
Keshavarz B, Peck K, Rezaei S et al (2022) Detecting and predicting visually induced motion sickness with physiological measures in combination with machine learning techniques. Int J Psychophysiol 176:14–26
Khoirunnisaa AZ, Pane ES, Wibawa AD et al (2018) Channel selection of EEG-based cybersickness recognition during playing video game using correlation feature selection (CFS). In: 2018 2nd international conference on biomedical engineering (IBIOMED). IEEE, pp 48–53
Kim YY, Kim HJ, Kim EN et al (2005) Characteristic changes in the physiological components of cybersickness. Psychophysiology 42(5):616–625
Kogler W, Wood G, Kober SE (2021) Effects of electrical brain stimulation on brain indices and presence experience in immersive, interactive virtual reality. Virtual Real, pp 1–11
Krokos E, Varshney A (2022) Quantifying VR cybersickness using EEG. Virtual Real 26(1):77–89
Kroupi E, Hanhart P, Lee JS et al (2014) User-independent classification of 2D versus 3D multimedia experiences through EEG and physiological signals. In: 8th International workshop on video processing and quality metrics for consumer electronics-VPQM 2014, CONF
Lawson B, Mead A (1998) The sopite syndrome revisited: drowsiness and mood changes during real or apparent motion. Acta Astronaut 43(3):181–192. https://doi.org/10.1016/S0094-5765(98)00153-2
Leuthardt EC, Schalk G, Wolpaw JR et al (2004) A brain–computer interface using electrocorticographic signals in humans. J Neural Eng 1(2):63
Li G, McGill M, Brewster S et al (2022) Multimodal biosensing for vestibular network-based cybersickness detection. IEEE J Biomed Health Inform 26(6):2469–2480. https://doi.org/10.1109/JBHI.2021.3134024
Liao CY, Tai SK, Chen RC et al (2020) Using EEG and deep learning to predict motion sickness under wearing a virtual reality device. IEEE Access 8:126784–126796
Lin CT, Chuang CH, Huang CS et al (2014) Wireless and wearable EEG system for evaluating driver vigilance. IEEE Trans Biomed Circuits Syst 8(2):165–176
Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 30:4765–4774
Lundberg SM, Nair B, Vavilala MS et al (2018) Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. Nat Biomed Eng 2(10):749
Maeda T, Ando H, Sugimoto M (2005) Virtual acceleration with galvanic vestibular stimulation in a virtual reality environment. In: IEEE proceedings. VR 2005. Virtual reality, 2005. IEEE, pp 289–290
Nam S, Jang KM, Kwon M et al (2022) Electroencephalogram microstates and functional connectivity of cybersickness. Front Hum Neurosci 16:857768
Oostenveld R, Praamstra P (2001) The five percent electrode system for high-resolution EEG and ERP measurements. Clin Neurophysiol 112(4):713–719
Parui S, Bajiya AKR, Samanta D et al (2019) Emotion recognition from EEG signal using XGBoost algorithm. In: 2019 IEEE 16th India council international conference (INDICON). IEEE, pp 1–4
Pedregosa F, Varoquaux G, Gramfort A et al (2011) Scikit-learn: machine learning in python. J Mach Learn Res 12:2825–2830
Pijeira-Díaz HJ, Drachsler H, Kirschner PA et al (2018) Profiling sympathetic arousal in a physics course: How active are students? J Comput Assist Learn 34(4):397–408
Plechawska-Wojcik M, Kaczorowska M, Zapala D (2019) The artifact subspace reconstruction (ASR) for EEG signal correction. A comparative study. In: Information systems architecture and technology: proceedings of 39th international conference on information systems architecture and technology–ISAT 2018: part II, Springer, pp 125–135
Poussot-Vassal C, Roy RN, Bovo A et al (2017) A loewner-based approach for the approximation of engagement-related neurophysiological features. In: International Federation of Automatic Control (IFAC) International conference proceedings

Prinzel LJ, Freeman FG, Scerbo MW et al (2000) A closed-loop system for examining psychophysiological measures for adaptive task allocation. Int J Aviat Psychol 10(4):393–410
Qu C, Che X, Ma S et al (2022) Bio-physiological-signals-based VR cybersickness detection. CCF Trans Pervasive Comput Interact 4(3):268–284
Raufi B, Longo L (2022) An evaluation of the EEG alpha-to-theta and theta-to-alpha band ratios as indexes of mental workload. Front Neuroinform 16:861967
Reason JT (1978) Motion sickness adaptation: a neural mismatch model. J R Soc Med 71(11):819–829
Rebenitsch L, Owen C (2016) Review on cybersickness in applications and visual displays. Virtual Real 20:101–125. https://doi.org/10.1007/s10055-016-0285-9
Recenti M, Ricciardi C, Aubonnet R et al (2021) Toward predicting motion sickness using virtual reality and a moving platform assessing brain, muscles, and heart signals. Front Bioeng Biotechnol 9:635661
Riccio GE, Stoffregen TA (1991) An ecological theory of motion sickness and postural instability. Ecol Psychol 3(3):195–240
Ronca V, Martinez-Levy AC, Vozzi A et al (2023) Wearable technologies for electrodermal and cardiac activity measurements: a comparison between fitbit sense, empatica E4 and shimmer GSR3+. Sensors 23(13):5847
Ryu K, Myung R (2005) Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. Int J Ind Ergon 35(11):991–1009
Shaffer F, Ginsberg JP (2017) An overview of heart rate variability metrics and norms. Front Public Health 5:290215
Sharples S, Cobb S, Moody A et al (2008) Virtual reality induced symptoms and effects (VRISE): comparison of head mounted display (HMD), desktop and projection display systems. Displays 29(2):58–69
Sheldon MR, Fillyaw MJ, Thompson WD (1996) The use and interpretation of the Friedman test in the analysis of ordinal-scale data in repeated measures designs. Physiother Res Int 1(4):221–228
Simon M, Schmidt EA, Kincses WE et al (2011) EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. Clin Neurophysiol 122(6):1168–1178
Stanney KM, Kennedy RS, Drexler JM (1997) Cybersickness is not simulator sickness. In: Proceedings of the human factors and ergonomics society annual meeting. SAGE Publications Sage CA, Los Angeles, CA, pp 1138–1142
Tanaka M, Shigihara Y, Ishii A et al (2012) Effect of mental fatigue on the central nervous system: an electroencephalography study. Behav Brain Funct 8(1):1–8
Tauscher JP, Witt A, Bosse S et al (2020) Exploring neural and peripheral physiological correlates of simulator sickness. Comput Anima Virtual Worlds 31:e1953. https://doi.org/10.1002/cav.1953
Teixeira J, Palmisano S (2021) Effects of dynamic field-of-view restriction on cybersickness and presence in HMD-based virtual reality. Virtual Real 25(2):433–445
Terpilowski MA (2019) scikit-posthocs: pairwise multiple comparison tests in python. J Open Sour Softw 4(36):1169
Tian N, Lopes P, Boulic R (2022) A review of cybersickness in head-mounted displays: raising attention to individual susceptibility. Virtual Real 26(4):1409–1441
Tietze H (2001) Stages of fatigue during long duration driving reflected in alpha related events in the EEH. In: International conference on traffic and transport psychology-ICTTP 2000, Held 4–7 September 2000, Berne, Switzerland-keynotes, symposia, thematic sessions, workshops, posters, list of participants and word viewer CD-ROM
Torres Vega M, Liaskos C, Abadal S et al (2020) Immersive interconnected virtual and augmented reality: a 5G and IoT perspective. J Netw Syst Manag 28:796–826
Treisman M (1977) Motion sickness: an evolutionary hypothesis. Science 197(4302):493–495
Van Gent P, Farah H, Nes N et al (2018) Heart rate analysis for human factors: development and validation of an open source toolkit for noisy naturalistic heart rate data. In: Proceedings of the 6th HUMANIST conference, pp 173–178
Virtanen P, Gommers R, Oliphant TE et al (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 17:261–272. https://doi.org/10.1038/s41592-019-0686-2
Warwick-Evans L, Symons N, Fitch T et al (1998) Evaluating sensory conflict and postural instability. Theories of motion sickness. Brain Res Bull 47(5):465–469. https://doi.org/10.1016/s0361-9230(98)00090-2
Wibirama S, Nugroho HA, Hamamoto K (2018) Depth gaze and ECG based frequency dynamics during motion sickness in stereoscopic 3D movie. Entertain Comput 26:117–127
Wolpaw JR, McFarland DJ, Neat GW et al (1991) An EEG-based brain–computer interface for cursor control. Electroencephalogr Clin Neurophysiol 78(3):252–259
Yang AHX, Kasabov NK, Cakmak YO (2023) Prediction and detection of virtual reality induced cybersickness: a spiking neural network approach using spatiotemporal EEG brain data and heart rate variability. Brain Inf 10(1):15
Yildirim C (2020) A review of deep learning approaches to EEG-based classification of cybersickness in virtual reality. In: 2020 IEEE international conference on artificial intelligence and virtual reality (AIVR), pp 351–357. https://doi.org/10.1109/AIVR50618.2020.00072
Young S, Adelstein B, Ellis S (2006) Demand characteristics of a questionnaire used to assess motion sickness in avirtual environment. In: IEEE virtual reality conference (VR 2006), pp 97–102. https://doi.org/10.1109/VR.2006.44
Zanto TP, Pan P, Liu H et al (2011) Age-related changes in orienting attention in time. J Neurosci 31(35):12461–12470
Zhai J, Barreto A, Chin C et al (2005) Realization of stress detection using psychophysiological signals for improvement of human–computer interactions. Proc IEEE SoutheastCon 2005:415–420. https://doi.org/10.1109/SECON.2005.1423280
Zhang C, Yu X (2010) Estimating mental fatigue based on electroencephalogram and heart rate variability. Pol J Med Phys Eng 16(2):67–84
Zhao C, Zhao M, Liu J et al (2012) Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator. Accid Anal Prev 45:83–90
Zheng WL, Lu BL (2017) A multimodal approach to estimating vigilance using EEG and forehead EOG. J Neural Eng 14(2):026017