Optimizing Visual Complexity for Physiologically-Adaptive VR Systems: Evaluating a Multimodal Dataset using EDA, ECG and EEG Features

Francesco Chiossi
LMU Munich
Germany
francesco.chiossi@um.ifi.lmu.de

Changkun Ou
LMU Munich
Germany
research@changkun.de

Sven Mayer
LMU Munich
Germany
info@sven-mayer.com

ABSTRACT

Physiologically-adaptive Virtual Reality systems dynamically adjust virtual content based on users' physiological signals to enhance interaction and achieve specific goals. However, as different users' cognitive states may underlie multivariate physiological patterns, adaptive systems necessitate a multimodal evaluation to investigate the relationship between input physiological features and target states for efficient user modeling. Here, we investigated a multimodal dataset (EEG, ECG, and EDA) while interacting with two different adaptive systems adjusting the environmental visual complexity based on EDA. Increased visual complexity led to increased alpha power and alpha-theta ratio, reflecting increased mental fatigue and workload. At the same time, EDA exhibited distinct dynamics with increased tonic and phasic components. Integrating multimodal physiological measures for adaptation evaluation enlarges our understanding of the impact of system adaptation on users' physiology and allows us to account for it and improve adaptive system design and optimization algorithms.

CCS CONCEPTS

• **Human-centered computing** → *User models.*

KEYWORDS

Physiological Computing, Virtual Reality, Adaptive Systems, Visual Complexity

ACM Reference Format:
Francesco Chiossi, Changkun Ou, and Sven Mayer. 2024. Optimizing Visual Complexity for Physiologically-Adaptive VR Systems: Evaluating a Multimodal Dataset using EDA, ECG and EEG Features. In *International Conference on Advanced Visual Interfaces 2024 (AVI 2024), June 03–07, 2024, Arenzano, Genoa, Italy*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3656650.3656657

1 INTRODUCTION

Physiological computing [15] is an emerging field that investigates how the physiological indicators of human affective and cognitive states can be utilized as inputs in adaptive systems to accomplish specific objectives, such as visualization, content or interaction adaptations [11]. Physiological computing adaptive systems are grounded in psychophysiological inference, which assumes that the measured physiological responses accurately and sensitively represent the underlying users' states [15]. However, this assumption often proves problematic due to the intricate relationship between physiological measures and their psychological or cognitive correlates [15]. The development of physiological computing is hindered by the complexity, validity, and specificity of psychophysiological inference [15]. Most physiological data exhibit complex dynamics, making it challenging to establish a one-to-one relationship with psychological constructs [15]. Given these challenges, it is essential to consider a multimodal evaluation of physiological interactions [16], validating multiple physiological data and their relationships for future use in adaptive systems.

Consequently, combining different physiological measurements enables a hybrid evaluation of adaptive systems [18]. Instead of solely measuring the final impact of an unsuccessful adaptation (e.g., a decrease in task performance), a second signal is utilized to assess the effect of the adaptation on users' arousal. This approach, known as multimodal or hybrid Brain-Computer Interface (BCI) systems, has been proposed to enhance the reliability, proficiency, and utility of BCI systems [18]. For example, Scherer et al. [42] combined a Heart Rate (HR) with Steady state visually evoked potential (SSVEP) BCI to enable self-initiation, allowing users to switch the adaptive system on/off independently. Thus, further research is needed to understand better the relationships between adaptive system dynamics, system performance, and its input physiological measures to effectively align them for user personalization and adaptation [38]. This is particularly relevant for adaptive systems, as multimodal input has been relatively overlooked or primarily focused on alternative channels for adaptation, such as peripheral measure or facial and gesture recognition [18, 38].

This work aims to evaluate the relationships between measures extracted from physiological signals, such as EDA, ECG, and EEG, and evaluate the impact of Virtual Reality (VR) system adaptation and visual complexity on these measures. This analysis is particularly important since different physiological signals may respond to adaptations at varying time windows, making some more suitable for faster-paced adaptations while others may require slower paces [15]. We chose the open dataset of Chiossi et al. [9], which includes a multimodal dataset (ECG, EDA, EEG, behavioral performance) collected during an interaction with two physiologically-adaptive systems that optimized VR visual distractors. We first investigated physiological reactivity in EDA, ECG, and EEG measures

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
AVI 2024, June 03–07, 2024, Arenzano, Genoa, Italy
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-1764-2/24/06
https://doi.org/10.1145/3656650.3656657

when the adaptive systems performed adaptations of visual complexity. The adaptive systems successfully impacted EDA features, specifically tonic and phasic components, in response to changes in visual complexity. This suggests that EDA can be a reliable indicator of cognitive workload and engagement, making it a promising candidate for adaptive system design. Secondly, analyzing stable visual complexity demonstrated selective influences on specific physiological features. While heart-related measures (HR and HRV) did not show significant effects, EEG measures (Alpha power and Alpha/Theta Ratio) exhibited strong linear associations with visual complexity. The increase in Alpha power and Alpha/Theta Ratio as visual complexity heightened suggested higher cognitive workload and engagement when participants saw visual distractors.

Our work extends the foundational work of Chiossi et al. [9] by investigating multimodal physiological responses to VR adaptations. While previous studies have established a baseline understanding of physiological signals as inputs for adaptive systems, our analysis advances this knowledge by evaluating how these signals vary over different adaptations and complexity levels.

2 RELATED WORK

We introduce physiological computing for VR and review adaptive systems that employ multimodal physiological inputs. Finally, we summarize the physiological correlates of visual complexity.

2.1 Physiological Computing in Virtual Reality

The physiological computing perspective allows the utilization of psychophysiological data for developing new control channels and adapting tasks based on changes in workload [36]. VR emerges as a promising domain [11], allowing for a high degree of flexibility and solution space that may not be feasible in physical spaces alone.

Physiologically adaptive VR systems leverage physiological input to discern users' states and adapt interaction features to achieve a shared goal. The combination of implicit physiological monitoring and adaptive VR environments establishes a closed-loop model [15]. Closed-loop models serve as a conceptual framework for adaptive systems that dynamically personalize software, visualizations, and interactions in real-time based on individual requirements [7]. Adaptive VR systems dynamically adjust system parameters in response to the current task or user. The adaptive controller embedded within a closed-loop system is a dynamic mechanism that responds to evolving inputs within predefined standards or goals. The primary objective has always been to enhance task performance, as exemplified by the seminal work of Pope and Bogart [36], where a biocybernetic loop was devised to sustain high engagement and optimize task performance. Similarly, Chiossi et al. [9] supported working memory (WM) performance by adapting the number of visual distractors with a rule-based approach, i.e., the level of visual complexity based on a tonic variation of EDA. Similar approaches have been proposed in applied training scenarios [30] and virtual rehabilitation therapies using fatigue-aware systems based on EMG [22]. Together, those results show how integrating physiological computing within adaptive VR systems can support task performance and improve the usability and user experience of VR systems and interactions.

2.2 Multimodal Physiological Inputs for Adaptive Systems

A multimodal adaptive system can be defined as a system that enables users to interact using two or more distinct modalities for both input and output [18]. In this context, "modality" encompasses the methods through which users provide input and the pathways through which they receive output information. An initial example system is by Pfurtscheller et al. [34], which employs two different brain signals (EEG and functional Magnetic Resonance Imaging) in the context of assistive technologies. Leeb et al. [22] combined EEG and electromyography (EMG) signals during left/right-hand movement tasks. To simulate muscle fatigue, they degraded the EMG signals and showed that the recognition performance could be improved compared to EEG or EMG-only recognition.

Furthermore, to enhance the adaptive system's performance, other research in the field has encompassed feature-level fusion [18]. This fusion technique is used for modalities that may not necessarily be of the same type but are tightly coupled or synchronized, such as ECG and EDA. Thus, such features are combined into a single vector as input into a classifier. For instance, Chanel et al. [6] employed both EEG and peripheral sensors for arousal assessment. They extracted six EEG and 18 peripheral features, merged them into a unified vector, and subsequently fed this data into classifiers such as naive Bayes and Fisher's discriminant [6]. Their results showed that sensor fusion yielded more robust results than utilizing EEG or peripheral physiological data in isolation.

Lastly, a third approach known as decision-level fusion has been utilized to integrate modalities that may not be tightly synchronized [47]. In this fusion level, data from each modality are independently modeled, and the individual recognition results are combined. Decision-level fusion allows for the combination of modalities to perform single or multiple tasks, contributing to a higher-level task. This type of fusion has been applied in various scenarios. For instance, Merzagora et al. [29] combined neurophysiological data, specifically EEG and fNIRS, in a WM task. However, it is worth noting that most of these studies have not extensively analyzed these features offline to evaluate the impact of adaptation on system usability and users' physiological reactions.

2.3 Physiological Correlates of Visual Complexity

The level of visual complexity, i.e., level of detail, clutter, and objects [31], in VR can influence cognitive load or attention allocation, ultimately affecting overall task performance and user experience [35, 37]. Thus, the limited capacity of perceptual and attentional processing in VR necessitates careful management of visual complexity to avoid cognitive overload, particularly in VR tasks requiring visual search [13] and executive functions like WM [28].

Furthermore, studies show that visual complexity affects physiological responses such as EDA and EEG. While moderate complexity enhances pleasantness and flow, facilitating cognitive engagement [12], excessive complexity can increase physiological stress and hinder performance [9, 35]. Peifer et al. [33] observed a U-shaped pattern in the psychophysiological processes of flow experience, where high flow values corresponded with moderate sympathetic and high parasympathetic activation. This aligns with

De Manzano et al. [12], suggesting a link between flow and increased physiological arousal, as measured by EDA or ECG. Such a balance of sympathetic and parasympathetic responses is associated with improved adaptation, effective workload coping [2], and enhanced stress management [41].

These results suggest that physiological arousal could mediate the relationship between visual complexity and flow. Moderate levels of complexity might induce an optimal level of physiological arousal, promoting flow experiences and enhancing task cognitive engagement. On the other hand, excessively high complexity could lead to increased physiological stress, potentially hindering flow experiences and task performance [24, 35].

Physiological arousal, as indicated by EDA, has been utilized to adjust VR visual complexity, enhancing task engagement and performance[9, 10]. Specifically, when adaptations of visual complexity in VR were task-relevant, they reduced users' perceived workload [10]. Conversely, task-irrelevant complexity improved performance and decreased workload [9] based on EDA. These findings highlight the potential of physiologically adaptive VR systems based on an adaptation of visual complexity to allow for more engaging and efficient experiences. Thus, in this work, we focus on the effect of VR adaptations on different physiological correlates of visual complexity to understand the user's real-time reactions to the adaptations made within the VR environment.

3 METHOD

We utilized the open dataset from Chiossi et al. [9] as it comprises behavioral, physiological (EEG, ECG, and EDA), and subjective data. Moreover, this dataset includes one physiological channel, i.e., EDA, as inputs for an adaptive system, thus together with EEG and ECG data allows for analyzing the effect of adaptations on such multimdoal data. The dataset is available on the Open Science Framework at https://osf.io/axvfy/. We refer the reader to their paper for a detailed task implementation and data collection description. The dataset included 20 participants with an average age of 26.05 years (SD = 3.62). EDA data were collected via a GSR module (250 Hz, BrainProducts GmbH, Germany). ECG data were recorded using a Polar H10 chest strap (130 Hz, Polar, Finland) with electrodes moistened before data collection and placed over the xiphoid process of the sternum. EEG data were recorded at a sampling rate of 250 Hz using a 7-channel dry electrode cap embedded in the HTC VIVE headset from Wearable Sensing (DSI-VR 300, 250 Hz, San Diego, CA, USA). The electrode positions followed the 10-20 system, including FCz, Pz, P3, P4, PO7, PO8, and Oz. Electrode impedances were maintained below 20 kΩ, with electrodes linked to the ears as a reference for the EEG recording. The EDA, ECG, and EEG data were simultaneously recorded using the LabStreamingLayer (LSL) framework[1].

3.1 Research Questions

We aim to gauge the users' responses to visual complexity adaptations and levels by assessing various physiological measures. This evaluation allows us to measure the impact of these changes on different physiological indicators. This will help to determine whether an adaptive system based on EDA can effectively include multiple modalities and be adapted for different applications. Moreover, we want to investigate the effect of different levels of visual complexity on attention allocation, engagement, and task load physiological correlates. Given the role of Skin Conductance Level (SCL), a tonic component of EDA, as input for adaptation, we hypothesize that system adaptations will significantly impact SCL (**HP1**). **HP1** is rooted in SCL's sensitivity to changes, a key indicator of the system's ability to respond to user physiological states dynamically. Such an effect would validate our system's architecture, highlighting the effectiveness of our physiologically adaptive VR system design. Our analysis contributes to developing physiologically-adaptive VR systems and drafts new possibilities for a larger input space and evaluation for immersive adaptive environments with varying levels of visual complexity. Drawing on the principles of physiologically-adaptive VR systems and existing research, we put forth the following hypothesis and research questions:

- **RQ1** Do adaptations of visual complexity that use EDA as input, impact participants' stress levels, as indexed by HR and HRV?
- **RQ2** Do adaptations of visual complexity that adjust the number of distractors, using EDA as input, influence internal and external attention, as measured by alpha and theta EEG oscillations?

Secondly, Chiossi et al. [9] provided an initial insight into the relationship between visual complexity, physiological arousal, and behavioral performance. Thus, we expand their work by evaluating the effect of varying stable levels of visual complexity on different physiological measures. Thus, we hypothesize:

- **RQ3** Does an increase in visual complexity, i.e., distracting information, increase external attention resources as indexed by a decrease in EEG Alpha oscillations [25]?
- **RQ4** Does an increase in visual complexity, i.e., distracting information, increase the cognitive workload as indexed by the ratio between alpha and theta oscillations [39]?

3.2 Experimental Task

The experimental task employed in this study was adapted from the N-Back task, as described by Chiossi et al. [9]. Participants were immersed in a neutral VR environment, where they were presented with a marble-like pillar and two buckets positioned on the left and right sides, respectively. Spheres of different colors (green, red, blue, and black) were generated and appeared on the pillar randomly, based on McMillan et al. [28]. Participants were required to use an HTC VIVE controller to grab the spheres and place them into the appropriate buckets. The rule for placing the spheres was based on matching the current sphere's color with the sphere's color presented two steps prior. The sphere would be placed in the right bucket if the colors matched. Conversely, if the colors did not match, the sphere would be placed in the left bucket. Participants had a time window of 4 seconds to pick up the sphere to avoid making an error. New spheres appeared when the current sphere was successfully placed in one of the buckets or after the 4-second time limit had elapsed.

[1] https://labstreaminglayer.readthedocs.io/info/intro.html

(a) Stream = 24 (b) Stream = 191 (c) Stream = 347

Figure 1: VR capture of a single trial of the VR n-back from a birds-eye perspective in the first row. In (a) is depicted the condition with low visual complexity with Stream = 24. In (b) is depicted the condition of moderate visual complexity with Stream = 191, and lastly in (c), the highest visual complexity with Stream = 347.

3.3 Adaptive Systems Architecture

Chiossi et al. [9] introduced an adaptive system that supports users' engagement and task performance by adapting the visual complexity based on EDA. Our focus was on the two adaptive conditions, where we divided the EDA, ECG, and EEG signals into 20-second epochs corresponding to the periods when the stream of non-player characters (NPCs) underwent adaptations, see Figure 1. The first adaptive system, i.e., Adaptive Test System is based on the Motivational Intensity model (MIM) [40]. According to this model, when task demands are perceived as achievable, there is a proportional relationship between mental effort and task demand [40]. However, as task demands increase and success becomes less likely, the investment of effort decreases while the perceived workload increases. This results in impaired performance, increased perceived workload, and reduced engagement, highlighting the association between WM capacity and these outcomes. Thus, the Adaptive Test System decreases the visual complexity when the arousal increases, as indexed by SCL, by removing 8 NPCs and increasing it by adding 16 NPCs when a decrease in SCL is detected. The second system, i.e., Reverse Adaptive System, followed an inverse logic and served as a control condition. Here, the system either aims to progressively increase the task demands by adding 16 NPCs when the participant's arousal is increasing or when the arousal is decreasing, removing visual complexity from the VR scene (-8 NPCs), ultimately leading to either an overly distracting or empty VR scene, to decrease users' engagement. The VR-physiologically adaptive system performed an average of $M = 6.94$ adaptations ($SD = 2.77$) for the Adaptive Test condition, while the Reverse Adaptive System $M = 5.19$ ($SD = 2.76$) adaptations.

3.3.1 Adaptive Mechanism. Both systems employ a rolling window approach for adaptation. They utilize two distinct data windows for SCL analysis: a 180-second window (W_1) for low-frequency changes and a 30-second window (W_2) for high-frequency changes. The SCL levels are averaged over these windows to stabilize the value, using an epsilon parameter for smoothing.

3.3.2 Slope Analysis. Slopes of SCL changes in these windows (s_1 and s_2) are computed, forming the basis for adaptive decision-making. s_1 is calculated from the average tonic value between points t_{-2} and t_0, while s_2 uses values between t_{-1} and t_0.

3.3.3 Rule-Based Adaptation. The adaptation decision is based on comparing the low-frequency slope (s_1) to the high-frequency slope (s_2), adhering to a threshold parameter θ to ensure stability in adaptations. This comparison drives the system to increase or decrease task difficulty, as detailed in Equation 1. This adaptive process occurs every 20 seconds, ensuring timely responsiveness to physiological changes.

$$adaptation(s_1, s_2) = \begin{cases} \texttt{increase} & \text{if } s_1 \leq s_2 - \theta \\ \texttt{decrease} & \text{if } s_1 \geq s_2 + \theta \end{cases} \tag{1}$$

4 DATA PREPROCESSING AND ANALYSIS

We investigated the physiological indicators of cognitive workload and arousal in a visual WM task, while adaptive systems dynamically adjusted the visual complexity. We open-source our analysis scripts on Github [2]. We invite researchers to reproduce our results and expand upon our findings and analysis approaches.

4.1 EDA & ECG Preprocessing

We used Neurokit [26] for EDA data preprocessing. This involved a third-order Butterworth high-pass filter at 3 Hz and nonnegative deconvolution analysis [3] to separate tonic and phasic components. We computed average amplitude of nsSCRs and tonic SCL, with nsSCR peaks identified using a .05μS threshold. ECG data were processed in the time domain, focusing on HR and HRV (RMSSD). We

[2] https://github.com/mimuc/avi24-adaptation-dataset

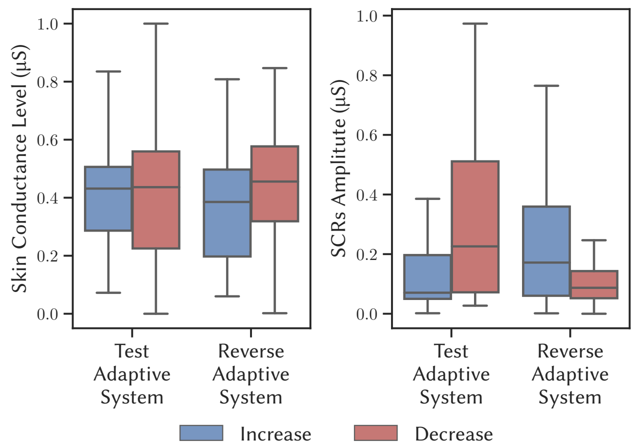

Figure 2: SCL and nSCRs Results. The *Adaptive System* significantly affected SCL, leading to distinct SCL responses based on stream direction. Amplitude of skin conductance responses (nSCRs) varies significantly with the adaptive system and adaptive visual complexity.

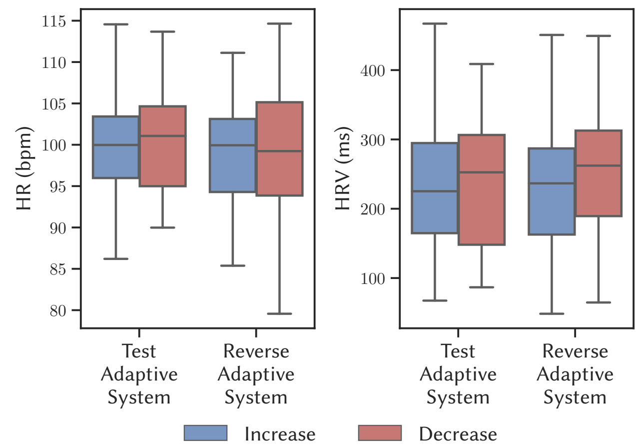

Figure 3: HR and HRV Results. We did not find significant results for variation of visual complexity in the Test and Reverse Adaptive Sytems in both HR and HRV.

applied a 3-45 Hz (3rd order) FIR band-pass filter with Neurokit [26] and used Hamilton's method for segmentation and QRS complex identification, to compute HR and HRV.

4.2 EEG Preprocessing

We processed the EEG raw data via the MNE Toolbox [17]. EEG data were recorded with a sampling frequency of 250 Hz from dry electrodes placed on Fz, P3, Pz,P4, PO7, Oz, PO8 locations (10/20 system), with a reference set at linked earlobes. We notch-filtered the signal at the power frequency of 50 Hz and then band-passed between 1 and 70 Hz to remove high and low-frequency drifts. We referenced the data to the common average reference (CAR). Next, we computed an independent component analysis (ICA) with extended infomax algorithgm for automatic artifact detection and correction with the ICLabel plugin [23]. We then analyzed the preprocessed EEG data in two frequency bands: Theta (4–8 Hz) and Alpha (8–12 Hz) using Welch's method. We computed alpha for posterior sites, i.e., PO8, PO7, and Oz electrodes, and extracted Theta power from midline sites, i.e. Fz and Pz. Moreover, we computed the ratio of midline theta activity's absolute power to posterior alpha activity's absolute power as an implicit measure of workload [39].

5 RESULTS

We report quantitative findings from analyzing physiological (EDA, ECG, and EEG) data from the dataset. We employed a Repeated Measures analysis of Variance [3] (RM-ANOVA) for adaptive levels of visual complexity and a Linear Mixed Model (LMM) approach for EDA, ECG, and EEG measures for stable visual complexity. To account for the repeated-measures structure in the data, we included a random intercept for each participant in our model.

[3] The predicted marginal means (PMM) for the different levels of the variable 'Adaptive System' and 'Stream Adaptation' were calculated using a Kenward-Roger degrees-of-freedom method.

5.1 Adaptive Visual Complexity

5.1.1 EDA Results – SCL. RM-ANOVA test detected a significant effect of the factor *Adaptive System* on the SCL ($F(1, 177) = 22.447$, $p < .001$). Tukey posthoc test showed that participants who experienced an Increase in the Test adaptive system condition had a significantly lower SCL compared to those in the Reverse adaptive system when experiencing a Stream Decrease ($PMM = -11506$, $SE = .458$, $p = .003$). When contrasting an Increase in the Test adaptive system with a Decrease in the same system, participants had a significantly higher SCL ($PMM = -1.515$, $SE = .41$, $p = .002$). Finally, when the Reverse adaptive system performed a Stream Increase, the SCL increased as compared to the Decrease in the same system ($PMM = -2.292$, $SE = .481$, $p < .001$). Similarly, in the pairwise comparison between a Decrease in the Test adaptive system and a Decrease in the Reverse Adaptive system, we report a significant difference ($PMM = -2.201$, $SE = .443$, $p < .001$). On the other hand, the factor *Stream Adaptation* did not show any effect ($F(1, 177) = 0.219$, $p = .640$).

5.1.2 EDA Results – nSCRs amplitude. RM-ANOVA on average nSCRs amplitude revealed a significant interaction effect between Test Adaptive and Reverse Adaptive systems and Stream Adaptation, $F(1, 177) = 36.09, p < .001$. However, the main effects of Adaptive System, $F(1, 177) = 0.86, p = .355$, and Stream Adaptation, $F(1, 177) = 1.52, p = .219$, were not statistically significant. Posthoc comparisons using the Tukey method revealed several significant differences. The contrast comparing the Stream Increase condition (Test adaptive system) to the Stream Increase condition (Reverse adaptive system) yielded a significant difference, with an estimated mean difference of 1.61 ($SE = 0.458$, $p = .003$). The contrast comparing the Stream Increase condition (Test adaptive system) to the Stream Decrease condition (Test adaptive system) we found a significant difference, with an estimated mean difference of 1.52 ($SE = .410$, $p = .002$). The contrast comparing the Stream Increase condition (Reverse adaptive system) to the

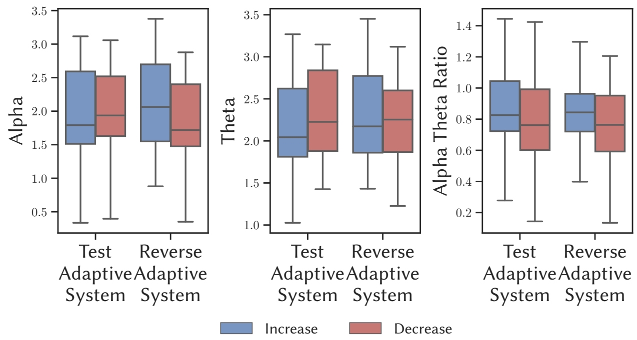

Figure 4: EEG results for Adaptive Visual Complexity. RM-ANOVA showed no significant effects for EEG Alpha power, neither from the ADAPTIVE SYSTEM nor STREAM ADAPTATION factors, and no significant interaction. Similar results were observed for Theta and Alpha to Theta Ratio.

STREAM DECREASE condition (REVERSE adaptive system) demonstrated a significant difference, with an estimated mean difference of -2.29 ($SE = 0.481$, $p < .001$). Likewise, the comparison comparing the STREAM DECREASE condition (TEST adaptive system) to the STREAM DECREASE condition (REVERSE adaptive system) revealed a significant difference, with an estimated mean difference of -2.20 ($SE = 0.443$, $p < .001$).

5.1.3 ECG Results – HR. The results of the HR RM-ANOVA analysis did not yield any significant findings, see Figure 3. As neither the ADAPTIVE SYSTEM factor ($F(1, 177) = 1.418$, $p = .235$), nor the STREAM ADAPTATION factor ($F(1, 177) = .136$, $p = .713$), or their interaction ($F(1, 177) = .055$, $p = .815$) had a significant effect on HR.

5.1.4 ECG Results – HRV. RM-ANOVA on HRV showed non-significant effects for the ADAPTIVE SYSTEM factor ($F(1, 177) = 0.338$, $p = .562$), the STREAM ADAPTATION factor ($F(1, 177) = .304$, $p = .582$), and their interaction ($F(1, 177) = .15$, $p = .699$), see Figure 3.

5.1.5 EEG Results – Alpha. RM-ANOVA for the EEG alpha power revealed non-significant effects. The ADAPTIVE SYSTEM factor ($F(1, 177) = .193$, $p = .661$) and the STREAM ADAPTATION factor ($F(1, 177) = .073$, $p = .787$) did not have a significant impact on alpha power. The interaction between the two main factors was also non-significant ($F(1, 177) = .223$, $p = .637$), see Figure 4.

5.1.6 EEG Results – Theta. Theta yielded similar results as Alpha. We did not detect significant effects across main factors ($p > .05$). See Figure 4.

5.1.7 EEG Results – Alpha / Theta Ratio. The EEG Alpha to Theta Ratio analysis yielded no significant effects ($p > .05$), see Figure 4.

5.2 Stable Visual Complexity

5.2.1 ECG Results – HR. We conducted a linear mixed model analysis to predict Heart Rate using Visual Complexity Level. We included participants as a random effect. The model's total explanatory power was substantial (conditional $R^2 = 0.78$). However, the effect of the ADAPTIVE SYSTEM was statistically non-significant and negative ($\beta = -0.14$, 95% CI [-0.51, 0.22], $p = .443$), see Figure 5.

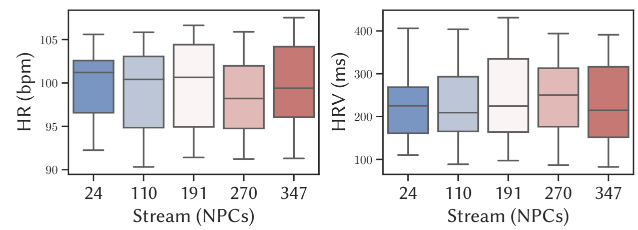

Figure 5: ECG HR and HRV results for Stable Visual Complexity. Analysis did showed no significant impact of the Visual Complexity on HR and HRV.

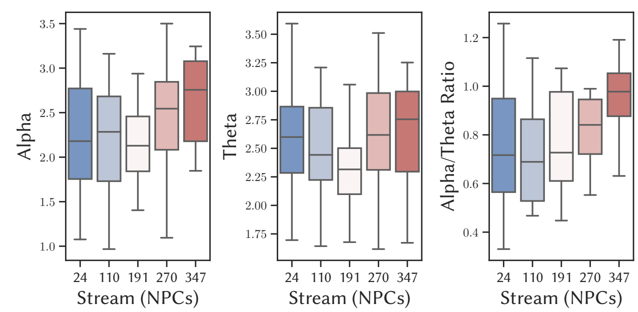

Figure 6: EEG results for Stable Visual Complexity. Analysis revealed that increasing Visual Complexity increases Alpha power and Alpha to-theta ratio while Theta power remains unaffected.

5.2.2 ECG Results – HRV. An LMM analysis was performed to predict HRV using Visual Complexity. As for HR, the model included participants as a random effect. The model exhibited substantial explanatory power (conditional R2 = .84). However, the effect of Visual Complexity was statistically non-significant and positive ($\beta = 1.22$, 95% CI [-4.59, 7.03], $p = .677$), see Figure 5.

5.2.3 EEG Results – Alpha. A linear mixed model analysis was performed to examine the relationship between EEG Alpha power and Visual Complexity. The model included participants as a random effect. The model exhibited a substantial total explanatory power (conditional R2 = .27), indicating its ability to explain the variability in the data. The effect of Visual Complexity on Alpha power was statistically significant and positive ($\beta = 1.12$, 95% CI [.14, 2.10], $p = .025$), suggesting that as Visual Complexity increases, there is a corresponding alpha synchronization, i.e., every increase in Visual Complexity level increases mean Alpha power by about 1.12 Hz.

5.2.4 EEG Results – Theta. In predicting Theta power as a function of Visual complexity, we found that model's total explanatory power was moderate ($R^2 = 0.26$). We did not report any significant effects effect ($\beta = .42$, 95% CI [-.53, 1.38], $p = .376$). The model intercept is 12.69 (95% CI [9.09, 16.29], $t(76) = 7.03$, $p < .001$).

5.2.5 EEG Results – Alpha / Theta Ratio. The model for prediction Alpha / Theta Ratio as a function of Visual Complexity showed a substantial explanatory power (R2 = .38) with moderate effect size (marginal R2 = .07). The intercept was estimated as 0.67 (95% CI [0.47, 0.87]). We found a statistically significant and positive effect of Visual Complexity (β = .07, 95% CI [.02, .12], p = .005).

6 DISCUSSION

We presented an in-depth analysis of the effect of visual complexity adaptation in VR based on physiological arousal on ECG, EEG, and EDA. Here, we discuss our results regarding the outcome of adaptive and stable levels of visual complexity. Finally, we discuss how such results can inform future applications of physiological computing and adaptive systems.

6.1 Adaptation of Visual Complexity

We interpret our results in light of the Motivational Intensity Model (MIM) [40], which provided the theoretical background to the adaptation logic for both systems [9]. The MIM provides a framework for understanding the relationship between task engagement and task demands on both behavioral and physiological levels.

According to the MIM, when task demands are manageable, and individuals confidently achieve successful performance, there is a proportional relationship between mental effort and task demand. However, as task demands increase and success becomes less likely, individuals reduce their effort investment, increasing perceived workload and decreasing engagement. Studies on EDA consistently demonstrated that increased physiological arousal is associated with higher task engagement and mental effort [16, 27].

Considering **HP1**, our findings indicate that the adaptive systems, specifically designed based on EDA, successfully influenced the EDA features in response to changes in visual complexity. This confirms the reliability and effectiveness of the adaptive systems in manipulating visual complexity by modulating EDA.

In the Test adaptive system, increased visual complexity led to higher arousal levels, promoting greater engagement. Conversely, increased visual complexity in the Reverse adaptive system resulted in decreased arousal levels, potentially leading to participant disengagement or boredom. Reducing visual complexity in the task-irrelevant elements of the Reverse system led to lower SCL, capturing the interference with performance and inducing a state of low arousal. This supports the notion that manipulating visual complexity impacts participants' physiological responses. Regarding **RQ1** and **RQ2**, we did not observe significant effects in the ECG and EEG measures during the 20-second window of visual complexity adaptations. This suggests that the impact of adaptations on these physiological measures may be impacted by trial-by-trial fluctuations or require long observation periods to have an effect.

First, it is possible that the 20-second window was not sufficient to detect subtle changes induced by visual complexity adaptations. Physiological responses, especially HR and HRV, may exhibit slower dynamics and require a longer period to discriminate across adaptations [5]. Secondly, visual complexity may not impact HR and HRV in general, as also shown by our results in Stable visual complexity.

Secondly, while these measures are sensitive indicators of physiological arousal and cardiac regulation, their response to visual complexity may vary depending on the specific context and application. Adaptive systems designed to target interventions related to anxiety disorders or stress inoculation may elicit more pronounced changes in HR and HRV, as they are specifically tailored to modulate physiological responses associated with these conditions [8, 19].

Regarding our **RQ3**, we found no effect of the visual complexity adaptations on EEG correlates of attention, task engagement, or mental workload. The 20-second intervals for the adaptations and the reliance on EDA variations may not have been sufficient to capture the rate of change of EEG measures related to attention, engagement, and workload. The not stationary nature of EEG signals [43], along with the specific frequencies and patterns associated with these cognitive processes, may require more refined adaptation mechanisms, different thresholds, and longer time windows to detect meaningful changes.

6.2 Stable Levels of Visual Complexity

Here, we investigated the relationship between a linear increase in visual complexity and its impact on computed physiological features related to physiological arousal, workload, and engagement.

6.2.1 ECG. Even though Chiossi et al. [9] found a relationship between visual complexity and physiological arousal as measured via EDA, we did not replicate such finding in arousal-related ECG measures. This could be attributed to several factors. First, the visual distractors used in our study were neutral and low-poly, which might not have sufficiently impacted the arousal state at a cardiac level. Prior research has indicated that visual stimuli with emotional content or higher arousing properties are more likely to induce significant changes in physiological measures such as HR and HRV [40]. Hence, additional manipulations may be needed to elicit stronger physiological responses to visual complexity, such as introducing emotionally charged visual distractors.

The effect of visual complexity on ECG may vary with task demands. Prior research indicates that task difficulty and cognitive load can alter how visual stimuli affect physiological responses[14, 32]. In our study, the cognitive requirements of the N-Back WM task might have overshadowed any influence of visual complexity on ECG, particularly considering the added demands of distractors on WM capacity.

Finally, the absence of significant effects on ECG measures by adaptations of visual complexity further supports the notion that the relationship between visual complexity and ECG responses is complex and context-dependent.

6.2.2 EEG. We investigated the effects of visual complexity on EEG correlates, specifically focusing on Alpha, Theta power, and the Alpha/Theta Ratio as indicators of attentional resources, engagement, and cognitive workload, respectively. Our hypotheses, **RQ3** and **RQ4**, proposed that increased visual complexity would lead to a decrease in Alpha oscillations and an increase in the Alpha/Theta Ratio. However, our results have different findings that do not confirm these hypotheses.

Contrary to previous work, we found a positive relationship between visual complexity and Alpha power. As visual complexity increased, so did Alpha power. This finding contradicts the notion that Alpha power reflects reduced attentional resources or external attentional engagement in response to visual complexity [44]. Instead, we propose two alternative interpretations that consider the potential role of mental fatigue or attentional withdrawal induced by continuous exposure to visually complex stimuli as in the stable visual complexity conditions.

Mental fatigue, linked to higher Alpha power, is thought to stem from cognitive resource depletion, as seen with continuous exposure to visually complex stimuli [44]. However, its mechanisms are debated, with some suggesting cognitive underload as a cause. Our study, indicating decreased accuracy with higher complexity, counters the underload hypothesis [44]. High Alpha power may reflect attentional withdrawal or task disengagement in demanding conditions [46], suggesting a shift to internal focus or reduced external attention. We interpret increased Alpha power as a compensatory response to preserve cognitive resources under mental fatigue [44].

We observed no significant modulation of frontal theta power with varying levels of visual complexity. This suggests a more complex and context-dependent relationship between visual complexity and frontal theta oscillations than previously thought [46]. The lack of significant frontal theta modulation might be attributed to the N-Back task's relatively low difficulty or the insufficiently distracting nature of visual distractors. If cognitive demands were low, additional cognitive control might not have been necessary, or if the distractors were not disruptive enough, they might not have significantly influenced frontal theta activity.

We confirmed **RQ5**, establishing a link between increased visual complexity and heightened EEG indicators of cognitive workload. The Alpha/Theta Ratio, a known marker of cognitive workload, linearly increased with visual complexity, indicating elevated cognitive demands during the N-Back task [31, 39]. This rise suggests higher cognitive resource and attentional control requirements to manage tasks amid increasing complexity [13]. The joint increase of Alpha and Theta powers reflects participants' efforts to cope with the task's demands and complex visual stimuli [39].

6.3 Insights for Physiologically-Adaptive System Design

Our analysis revealed increased Alpha power with greater visual complexity, challenging the view that Alpha power signifies reduced attention in complex visual scenarios [25]. This increase might indicate mental fatigue due to continuous exposure to complex stimuli [44]. In physiologically adaptive systems, monitoring Alpha power can be crucial for detecting mental fatigue, which may lead to distraction and lower cognitive performance [4]. By adapting to detected mental fatigue, such as adjusting task demands, these systems can maintain user performance and engagement [15].

In high-stakes training scenarios, such as medical simulations or hazardous environment training, users must maintain optimal cognitive performance for effective learning and decision-making [1]. Adaptive systems, using continuous Alpha power monitoring, can detect early signs of mental fatigue and adjust training complexity or introduce breaks [45]. Similarly, in MR collaborative workspaces, these systems can manage visual information to prevent cognitive overload [21]. The increased Alpha/Theta Ratio, indicative of higher workload in complex tasks, supports using this metric in adaptive systems for dynamic task complexity adjustment [39].

In adaptive Mixed Reality (MR) environments, particularly in transitional interfaces, cognitive demand varies during transitions across the MR continuum [20]. Adaptive systems can use the Alpha/Theta Ratio to adjust visual complexity, aiding smoother transitions between VR, AR, and physical reality. For example, reducing complexity in VR during high cognitive workload or simplifying visual elements in AR during mental fatigue enhances user adaptation. Based on real-time Alpha/Theta Ratio input, these adaptations can prevent cognitive overload and improve engagement. Here, HR and HRV, unaffected by visual complexity, may not assess cognitive demands in complex tasks but are promising in affective computing applications [30].

6.4 Limitation and Future Work

Our study links visual complexity with physiological responses, underscoring the potential of adaptive systems in VR and MR, while identifying limitations and space for further research. Our analysis focused on investigating statistical differences in physiological responses to visual complexity. Future work should explore classification approaches using machine learning algorithms to enhance our findings' accuracy and precision [18]. Utilizing classifiers allows a more efficient understanding of the relationship between physiological measures and visual complexity, on the amount of data needed for accurate classification and hardware performance threshold. We propose employing supervised transfer learning or unsupervised self-correcting classifiers, which require minimal explicit training phases. This approach can improve the robustness of the results generalizing to diverse task set.

Moreover, we did not investigate change in adaptation quality over time. User experience and usability may evolve with prolonged exposure to adaptive systems [43]. Factors such as learning curves, habituation, and system predictability can significantly influence users' perceptions and interactions with the system. To address this limitation, future studies should conduct longitudinal experiments with multiple sessions per participant to capture the dynamics of adaptation and user experience over time [43].

7 CONCLUSION

We investigated the effect of visual complexity levels and adaptation over physiological correlates of workload, engagement, and attention allocation. Notably, EDA demonstrated significant reactivity to visual complexity adaptations, suggesting its reliability as an indicator of cognitive workload and engagement. We emphasize the importance of multimodal evaluation of physiological interactions to understand the relationships between physiological responses and users' physiological reactivity when interacting with an adaptive system. Integrating multiple physiological measures and employing them as an evaluation metric and multimodal input can significantly enhance the effectiveness of adaptive systems, aligning them more effectively with users' complex interactions.

ACKNOWLEDGMENTS

Francesco Chiossi and Sven Mayer were supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project ID 251654672 TRR 161.

REFERENCES

[1] Laurence Alison, Claudia Van Den Heuvel, Sara Waring, Nicola Power, Amy Long, Terence O'Hara, and Jonathan Crego. 2013. Immersive simulated learning environments for researching critical incidents: A knowledge synthesis of the literature and experiences of studying high-risk strategic decision making. *Journal of Cognitive Engineering and decision making*.

[2] Richard W Backs, John K Lenneman, and Jamie L Sicard. 1999. The Use of Autonomic Components to Improve Cardiovascular Assessment of Mental Workload in Flight.. *The Int. J. of Aviation Psychology*.

[3] Mathias Benedek and Christian Kaernbach. 2010. Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, 647–658.

[4] Kyle A. Bernhardt, Dmitri Poltavski, Thomas Petros, and F. Richard Ferraro. 2019. Differentiating Active and Passive Fatigue with the use of Electroencephalography. *Proc. of HFES.*

[5] Nicolas Bourdillon, Laurent Schmitt, Sasan Yazdani, Jean-Marc Vesin, and Grégoire P Millet. 2017. Minimal window duration for accurate HRV recording in athletes. *Front. in neuroscience.*

[6] Guillaume Chanel, Julien Kronegg, Didier Grandjean, and Thierry Pun. 2006. Emotion assessment: Arousal evaluation using EEG's and peripheral physiological signals. In *Int. workshop on multimedia content representation*. Springer.

[7] Francesco Chiossi, Changkun Ou, Carolina Gerhardt, Felix Putze, and Sven Mayer. 2023. Designing and Evaluating an Adaptive Virtual Reality System using EEG Frequencies to Balance Internal and External Attention States. *arXiv preprint arXiv:2311.10447*.

[8] Francesco Chiossi, Changkun Ou, and Sven Mayer. 2023. Exploring Physiological Correlates of Visual Complexity Adaptation: Insights from EDA, ECG, and EEG Data for Adaptation Evaluation in VR Adaptive Systems. In *Proc. CHI EA*. ACM.

[9] Francesco Chiossi, Yagiz Turgut, Robin Welsch, and Sven Mayer. 2023. Adapting Visual Complexity Based on Electrodermal Activity Improves Working Memory Performance in Virtual Reality. *Proc. ACM Hum.-Comput. Interact.*

[10] Francesco Chiossi, Robin Welsch, Steeven Villa, Lewis Chuang, and Sven Mayer. 2022. Virtual Reality Adaptation Using Electrodermal Activity to Support the User Experience. *Big Data and Cognitive Computing*.

[11] Francesco Chiossi, Johannes Zagermann, Jakob Karolus, Nils Rodrigues, Priscilla Balestrucci, Daniel Weiskopf, Benedikt Ehinger, Tiare Feuchtner, Harald Reiterer, Lewis L. Chuang, Marc Ernst, Andreas Bulling, Sven Mayer, and Albrecht Schmidt. 2022. Adapting visualizations and Interfaces to the User. *it-Information Technology*.

[12] Örjan De Manzano, Töres Theorell, László Harmat, and Fredrik Ullén. 2010. The psychophysiology of flow during piano playing. *Emotion*.

[13] John Duncan and Glyn W Humphreys. 1989. Visual search and stimulus similarity. *Psychological review*.

[14] Caroline Dussault, Jean-Claude Jouanin, Matthieu Philippe, and Charles-Yannick Guezennec. 2005. EEG and ECG changes during simulator operation reflect mental workload and vigilance. *Aviation, space, and environmental medicine*.

[15] Stephen H. Fairclough. 2009. Fundamentals of physiological computing. *Interacting with Computers*.

[16] Stephen H Fairclough and Louise Venables. 2006. Prediction of subjective states from psychophysiology: A multivariate approach. *Biological psychology*.

[17] Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, and Lauri Parkkonen. 2013. MEG and EEG data analysis with MNE-Python. *Front. in neuroscience*.

[18] Hayrettin Gürkök and Anton Nijholt. 2012. Brain–computer interfaces for multimodal interaction: a survey and principles. *Int. J. of HCI*.

[19] Linda Hirsch, Florian Müller, Francesco Chiossi, Theodor Benga, and Andreas Martin Butz. 2023. My Heart Will Go On: Implicitly Increasing Social Connectedness by Visualizing Asynchronous Players' Heartbeats in VR Games. *Proceedings of the ACM on Human-Computer Interaction*, 976–1001.

[20] Hans-Christian Jetter, Jan-Henrik Schröder, Jan Gugenheimer, Mark Billinghurst, Christoph Anthes, Mohamed Khamis, and Tiare Feuchtner. 2021. Transitional Interfaces in Mixed and Cross-Reality: A New Frontier?. In *Proc. of ISS (ISS '21)*. ACM.

[21] Lik-Hang Lee, Tristan Braud, Pengyuan Zhou, Lin Wang, Dianlei Xu, Zijun Lin, Abhishek Kumar, Carlos Bermejo, and Pan Hui. 2021. All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda.

[22] Robert Leeb, Hesam Sagha, Ricardo Chavarriaga, and José del R Millán. 2011. A hybrid brain–computer interface based on the fusion of electroencephalographic and electromyographic activities. *Journal of neural engineering*.

[23] Adam Li, Jacob Feitelberg, Anand Prakash Saini, Richard Höchenberger, and Mathieu Scheltienne. 2022. MNE-ICALabel: Automatically annotating ICA components with ICLabel in Python. *Journal of Open Source Software*.

[24] Christopher R Madan, Janine Bayer, Matthias Gamer, Tina B Lonsdorf, and Tobias Sommer. 2018. Visual complexity and affect: Ratings reflect more than meets the eye. *Front. in psychology*.

[25] Elisa Magosso, Francesca De Crescenzio, Giulia Ricci, Sergio Piastra, and Mauro Ursino. 2019. EEG alpha power is modulated by attentional changes during cognitive tasks and virtual reality immersion. *Computational intelligence and neuroscience*.

[26] Dominique Makowski, Tam Pham, Zen J. Lau, Jan C. Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and SH Chen. 2021. NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behav. research methods*.

[27] Matteo Marucci, Gianluca Di Flumeri, Gianluca Borghini, Nicolina Sciaraffa, Michele Scandola, Enea Francesco Pavone, Fabio Babiloni, Viviana Betti, and Pietro Aricò. 2021. The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence. *Scientific Reports*.

[28] Kathryn M McMillan, Angela R Laird, Suzanne T Witt, and M Elizabeth Meyerand. 2007. Self-paced working memory: Validation of verbal variations of the n-back paradigm. *Brain research*.

[29] Anna C Merzagora, Meltem Izzetoglu, Robi Polikar, Valerie Weisser, Banu Onaral, and Maria T Schultheis. 2009. Functional near-infrared spectroscopy and electroencephalography: a multimodal imaging approach. In *HCI International*. Springer.

[30] John E. Muñoz, M. Cameirão, S. Bermúdez i Badia, and E. Rubio Gouveia. 2018. Closing the Loop in Exergaming - Health Benefits of Biocybernetic Adaptation in Senior Adults. In *Proc. of CHI Play*. ACM.

[31] Aude Olivia, Michael L Mack, Mochan Shrestha, and Angela Peeper. 2004. Identifying the perceptual dimensions of visual complexity of scenes. In *Proc. of CCC*.

[32] Mark Parent, Vsevolod Peysakhovich, Kevin Mandrick, Sébastien Tremblay, and Mickaël Causse. 2019. The diagnosticity of psychophysiological signatures: Can we disentangle mental workload from acute stress with ECG and fNIRS? *Int. J. of Psychophysiology*.

[33] Corinna Peifer, André Schulz, Hartmut Schächinger, Nicola Baumann, and Conny H Antoni. 2014. The relation of flow-experience and physiological arousal under stress—can u shape it? *Journal of Experimental Social Psychology*.

[34] Gert Pfurtscheller, Brendan Z Allison, Günther Bauernfeind, Clemens Brunner, Teodoro Solis Escalante, Reinhold Scherer, Thorsten O Zander, Gernot Mueller-Putz, Christa Neuper, and Niels Birbaumer. 2010. The hybrid BCI. *Front. in neuroscience*.

[35] Andrea C Pierno, Andrea Caria, Scott Glover, and Umberto Castiello. 2005. Effects of increasing visual load on aurally and visually guided target acquisition in a virtual environment. *Applied ergonomics*.

[36] Alan T Pope, Edward H Bogart, and Debbie S Bartolome. 1995. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological psychology*.

[37] Eric D Ragan, Doug A Bowman, Regis Kopper, Cheryl Stinson, Siroberto Scerbo, and Ryan P McMahan. 2015. Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. *IEEE TVCG*.

[38] Rahul Rajan, Ted Selker, and Ian Lane. 2016. Task Load Estimation and Mediation Using Psycho-Physiological Measures. In *Proc. of IUI '16*. ACM.

[39] Bujar Raufi and Luca Longo. 2022. An Evaluation of the EEG alpha-to-theta and theta-to-alpha band Ratios as Indexes of Mental Workload. *Front. in Neuroinformatics*.

[40] M. Richter, G.H.E. Gendolla, and R.A. Wright. 2016. Chapter Five - Three Decades of Research on Motivational Intensity Theory: What We Have Learned About Effort and What We Still Don't Know. *Advances in Motivation Science*.

[41] Hartmut Schächinger, Johannes Port, Stuart Brody, Lilly Linder, Frank H Wilhelm, Peter R Huber, Daniel Cox, and Ulrich Keller. 2004. Increased high-frequency heart rate variability during insulin-induced hypoglycaemia in healthy humans. *Clinical Science*.

[42] Reinhold Scherer, GR Müller-Putz, and Gert Pfurtscheller. 2007. Self-initiation of EEG-based brain–computer communication using the heart rate response. *Journal of Neural Engineering*.

[43] Pradeep Shenoy, Matthias Krauledat, Benjamin Blankertz, Rajesh PN Rao, and Klaus-Robert Müller. 2006. Towards adaptive classification for BCI. *Journal of neural engineering*.

[44] Yvonne Tran, Ashley Craig, Rachel Craig, Rifai Chai, and Hung Nguyen. 2020. The influence of mental fatigue on brain activity: Evidence from a systematic review with meta-analyses. *Psychophysiology*, e13554.

[45] Leonard J Trejo, Karla Kubitz, Roman Rosipal, Rebekah L Kochavi, Leslie D Montgomery, et al. 2015. EEG-based estimation and classification of mental fatigue. *Psychology*.

[46] Edmund Wascher, Björn Rasch, Jessica Sänger, Sven Hoffmann, Daniel Schneider, Gerhard Rinkenauer, Herbert Heuer, and Ingmar Gutberlet. 2014. Frontal theta activity reflects distinct aspects of mental fatigue. *Biological psychology*.

[47] David Zhang, Fengxi Song, Yong Xu, and Zhizhen Liang. 2009. Decision level fusion. In *Advanced pattern recognition technologies with applications to biometrics*. IGI Global.

View publication stats