# The Butterfly Effect: Novel Opportunities for Steady-State Visually-Evoked Potential Stimuli in Virtual Reality

Jonas Auda
University of Duisburg-Essen
Essen, Germany
jonas.auda@uni-due.de

Uwe Gruenefeld
University of Duisburg-Essen
Essen, Germany
uwe.gruenefeld@uni-due.de

Thomas Kosch
TU Darmstadt
Darmstadt, Germany
kosch@tk.tu-darmstadt.de

Stefan Schneegass
University of Duisburg-Essen
Essen, Germany
stefan.schneegass@uni-due.de

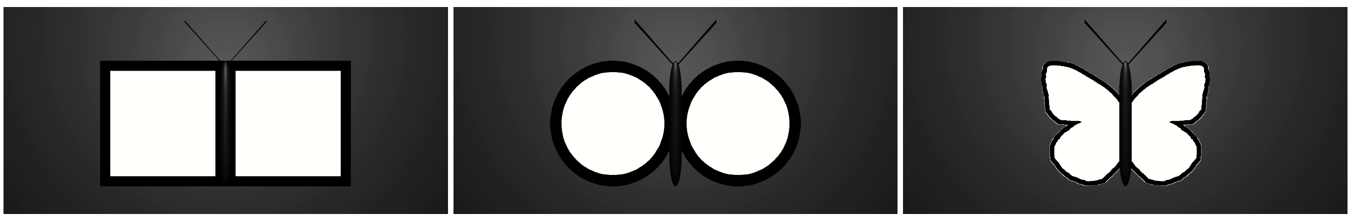

**Figure 1: Three butterflies with different levels of shape realism. Left: Low shape realism. Middle: Moderate shape realism. Right: High shape realism.**

## ABSTRACT

In Virtual Reality (VR), Steady-State-Visual Evoked Potentials (SSVEPs) can be used to interact with the virtual environment using brain signals. However, the design of SSVEP-eliciting stimuli often does not match the virtual environment, and thus, disrupts the virtual experience. In this paper, we investigate stimulus designs with varying suitability to blend in virtual environments. Therefore, we created differently-shaped, virtual butterflies. The shapes vary from rectangular wings, over round wings, to a wing shape of a real butterfly. These butterflies elicit SSVEP responses through different animations – flickering or flapping wings. To evaluate our stimuli, we first extracted suitable frequencies for SSVEP responses from the literature. In a first study, we determined three frequencies yielding the best detection accuracy in VR. We used these frequencies in a second study to analyze detection accuracy and appearance ratings using our stimuli designs. Our work contributes insights into the design of SSVEP stimuli that blend into virtual environments and still elicit SSVEP responses.

## CCS CONCEPTS

• **Human-centered computing → Virtual reality**; *User studies.*

## KEYWORDS

Virtual Reality, Brain-Computer Interface, Steady State Visually Evoked Potential, Realism

**ACM Reference Format:**
Jonas Auda, Uwe Gruenefeld, Thomas Kosch, and Stefan Schneegass. 2022. The Butterfly Effect: Novel Opportunities for Steady-State Visually-Evoked Potential Stimuli in Virtual Reality. In *Augmented Humans 2022 (AHs 2022), March 13–15, 2022, Kashiwa, Chiba, Japan.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3519391.3519397


*AHs 2022, March 13–15, 2022, Kashiwa, Chiba, Japan*

ACM ISBN 978-1-4503-9632-5/22/03…$15.00
https://doi.org/10.1145/3519391.3519397

## 1 INTRODUCTION

Steady-State Visual Evoked Potentials (SSVEPs) are cortical responses that occur when one is stimulated visually at a consistent frequency [56]. For example, viewing a light source at a constant frequency causes a measurable resonance at electrodes placed on the occipital lobe [11], a posterior part of the brain that is responsible for visual perception. Early work in neuroscience realized the potential of SSVEP, requiring only a low amount of data and short training times to achieve satisfactory results [55, 66]. Since cortical activity is passively generated, the use of SSVEP for people with physical impairments moved into the focus of research [12].

SSVEPs have become attractive in the domain of Human-Computer Interaction (HCI) due to their robustness [11], reasonable signal-to-noise ratio [48], and high input resolution [52], allowing one to reliably distinguish between light sources flickering at different frequencies using cortical activity. By looking at the respective flickering items, users can reliably select elements and type or read text [1, 22, 25], control hardware (e.g., wheelchairs [33]), or navigate virtual environments [24]. The visual presentation of SSVEP stimuli went through an evolution to better integrate them with current user interfaces. This ranged from the use of LEDs [61], abstract flickering elements on the computer screen [66], to the integration of flickering elements in mixed reality environments [31]. SSVEPs have been successfully evaluated in augmented reality [59] and are slowly becoming popular for immersive hands-free interaction in virtual reality (VR) [2]. VR benefits from the implicit integration of SSVEP stimuli into their environments, providing user experience designers with an additional interaction channel and a method to estimate which elements attract the user's attention.

In general, VR aims for immersive experiences that let users dive into a digital world that feels like their own reality. However, current representations of SSVEP stimuli are rendered in an abstract form of mostly flickering squares or circles, making it difficult to integrate SSVEP seamlessly into the virtual environment. Although past research reported that integrating SSVEP in VR can be successfully achieved [24], virtual experiences are disrupted by presenting abstract SSVEP stimuli which do not fit the virtual environment. We argue that SSVEP stimuli should blend into the environment, enabling a subtle interaction in VR. Past research investigated alternative SSVEP representations such as rotating items [43] or flickering menu navigation elements [2]. Hence, there exists a research gap regarding the comparison of the classification accuracy of abstract stimuli like menu items and blending objects that match the appearance of the VR environment. This poses the question: How does the SSVEP stimuli appearance affect the classification accuracy in VR?

In this work, we evaluate how different SSVEP parameters impact classification accuracy in VR. We start by surveying past HCI literature regarding commonly used SSVEP frequencies. We then performed our first study to compare these frequencies regarding their classification accuracy. We selected a triple of frequencies that yielded the highest classification accuracy for our second study. For this study, we designed SSVEP stimuli in form of butterflies with three levels of shape realism (*low*; *moderate*; *high*, see Figure 1). These stimuli elicit SSVEP responses through different animations –*flickering* wings or *flapping* wings.

In terms of classification accuracy, we found that *flickering* wings outperformed *flapping* wings. Our results show that the butterfly with quadratic and *flickering* wings yielded the highest accuracy (78.7%) for the classification of the three frequencies. For a *flapping* butterfly with real wing contours, we obtained a smaller classification accuracy (67.5%). Subjectively, participants perceived the butterfly with realistically-shaped wings as most natural in terms of appearance and movement. Our findings provide VR developers insights that help them better understand the trade-off between classification accuracy and a realistic appearance of SSVEP stimuli that blend with VR environments. We choose butterflies for our approach because the large wings size allowed us to use their anatomic properties to excite the VR users' retinas. Further, natural scenes are common in VR, therefore, we opted for a commonly occurring animal that has wings and can float in the air to make its appearance in front of the user plausible. Nonetheless, butterflies do not occur in every VR scenario. Therefore we used them as an example for SSVEP-based interaction in VR. To put our results into perspective, we outline more SSVEP stimuli integration opportunities as future work. Our work serves as an initial example for integrated SSVEP stimuli in VR to kick-off research that explores the underlying design space.

The contribution of our work is twofold: (1) We summarize commonly used SSVEP frequencies, which we evaluated in a first study (N=12) by utilizing machine learning to compare classification accuracies. (2) We performed a second study (N=12) evaluating how different realism levels of our SSVEP stimuli affect the classification accuracy and the associated subjective appearance ratings.

## 2 RELATED WORK

Previous research showed that SSVEPs have a high robustness, satisfactory information transfer rate, and good signal-to-noise-ratio [28], making them suitable for BCI-based interaction. Exemplary scenarios range from steering wheelchairs [29, 33], controlling prosthetic hands [35], providing text input through spellers [28, 57, 62], or interaction in VR [2, 4, 24, 50]. Available low-cost BCIs, such as the OpenBCI *Ganglion*[1] can be efficiently deployed as they do not require a rigorous setup of hardware [3]. As a result, a wide array of research emerged in HCI using the SSVEP interaction paradigm.

### 2.1 SSVEP Stimuli Characteristics

SSVEP stimuli can be displayed with different characteristics. Previous research investigated the robustness of SSVEP stimuli properties including different frequencies, sizes, shapes, colors, and patterns. SSVEP frequencies can be divided into three frequency ranges. Low frequencies are centered around $15Hz$, medium frequencies around $31Hz$, and high frequencies around $41Hz$ [60]. In this context, Kuś et al. identified a continuous range of suitable frequencies for strong SSVEP responses, ranging from $12Hz$ to $18Hz$ [26]. The selection of suitable frequencies is important since the frequency selection has an essential influence on the interaction performance [57].

SSVEP responses can be elicited through stimuli using blinking LEDs or flickering graphical elements on displays [66]. Previous approaches used differently colored LEDs, rendered black and white flickering shapes on displays, or pattern reversal stimuli which are alternating graphical patterns (e.g., checkerboards) [66] or motion-reversal stimuli [65]. Moreover, stimuli in motion like spinning [43] or repeatedly size-changing shapes [7] elicit SSVEP responses. Lately, SSVEP was adopted for interaction in augmented reality (AR) [58, 59] and VR [2, 32].

### 2.2 Interacting with SSVEP

SSVEP stimuli induce mental load and visual fatigue during interaction [36, 37]. Xie et al. compared periodic flickering to motion-reversal stimuli [64]. In terms of mental workload, motion reversal stimuli outperformed periodic stimuli. Long-term BCI interaction can also employ motion-reversal stimuli to reduce visual fatigue. To enhance the visual comfort of SSVEP stimuli, Rekrut et al. compared SSVEP responses of spinning icons to traditional flickering stimuli [43]. They showed that spinning icons could perform equally well than traditional SSVEP stimuli in terms of classification accuracy while they were perceived as less tiring. Many SSVEP-based approaches are restricted to abstract tasks. Therefore, everyday use cases were investigated by Bi et al. [5]. They evaluated SSVEP-based interaction with a heads-up display (HUD) integrated into a

[1]Ganglion Board, https://shop.openbci.com/collections/frontpage/products/ganglion-board, last checked on March 11, 2022.

vehicle's windshield. They classified SSVEP responses to control a simulated car. They showed that SSVEP-based HUD can indeed be used to control a car.

However, previous approaches rely mainly on 2D displays or LEDs [62] to elicit SSVEP responses. Other approaches used SSVEP in AR [59]. For example, controlling a smart home via a SSVEP-based AR- user interface (UI) [58]. In contrast to these approaches, VR relies heavily on interaction with a synthetic, 3D world. Previous approaches investigated a variety of SSVEP stimuli in immersive VR environments. Simply adopting 2D stimuli useful on flat displays in VR might not fit the three-dimensional character of virtual environments. For example, a black and white flickering square floating in mid-air in a VR fantasy game. Prior work showed that SSVEP in VR improves user engagement (e.g., higher information transfer rate) compared to 2D displays [24]. Hence, SSVEP in VR promises a wide array of interaction possibilities in the future. We are confident that the full potential of SSVEP-based interaction in VR does not rely on the classification accuracy only. Instead, we argue that the appearance of the presented stimuli should blend into the VR world rather than disrupting the virtual experience. To investigate this, we designed SSVEP stimuli that have the potential to blend with the VR environment while eliciting SSVEP responses that can be measured robustly. Before we introduce our approach, we introduce prior work to ground our research.

## 2.3 SSVEP-based Interaction in Virtual Reality

Choi et al. investigated classification accuracy and visual comfort of SSVEP stimuli in VR [7]. To control an avatar in a virtual environment, they employed two stimuli types – a grow shrink stimulus (GSS) and a pattern-reversal checkerboard stimulus (PRCS).

Stimuli that were subjectively more comfortable to participants showed higher classification accuracy. Nonetheless, the authors state that more investigations are needed to generalize the results. With this in mind, we designed our own set of SSVEP stimuli for VR with varying visual appearances.

Stawicki et al. employed an SSVEP-based virtual control of a vacuum cleaner robot in VR [50]. They compared traditional SSVEP stimuli on a 2D display to an immersive VR scenario. They achieved a better information transfer rate in VR as well as a lower task completion time than on a traditional PC setup. To navigate a virtual environment, Stawicki et al. compared SSVEP-based interaction with traditional PC environments to immersive VR using head-mounted display (HMDs) [49]. Through flickering rectangles, participants could move through the virtual environment by focusing on rectangles each associated with a specific movement. They found that in VR participants needed fewer commands to navigate the virtual environment. Further, participants traversed the environment 50% faster when using a VR-HMD. On top of that, participants were more aware of the virtual environment when using HMDs compared to a traditional PC setup. Ma et al. combined SSVEP-based BCI with eye tracking for text entry in VR [32]. Through the combination of these two modalities, they achieved a higher information transfer rate compared when using a single modality. They were able to achieve an input speed of 10 words per minute. In terms of accuracy, their VR approach outperformed similar approaches that relied on displays to elicit SSVEP responses [51]. A playful approach by Koo et al. used SSVEPs to move a ball through a maze in VR [24]. The ball was viewed from a bird-eye view. Around the ball, there were flickering squares. Focusing on one of the squares lets the ball move in the direction of the square. The goal was to steer the ball to the end of a maze. Through a study they found that interacting in VR resulted in shorter playtime and therefore higher information transfer rate than playing the game using a traditional 2D display.

In essence, previous research shows the great potential of SSVEP-based interaction in VR. In contrast, our work provides insights into the trade-off between satisfactory interaction stability and visual comfort for SSVEP stimuli in VR. We envision that a robust SSVEP-based contactless interaction for persons with physical impairments [33] can make future VR apps more accessible.

A closely related previous example for this is *Sublime*. Here, Armengol-Urpi et al. proposed a concept that incorporates stimuli into a VR environment [2]. In a virtual environment, users could focus on flickering movie covers in a virtual menu to select a movie to watch. While focusing on the movie covers, loading bars indicated when the selection is triggered. Then the movie was started. During playback, an additional object could be focused to get back to the previous menu. In this approach Armengol-Urpi et al. used higher frequencies (above $41Hz$ [60]). Similar to this approach, we integrated SSVEP stimuli into VR objects rather than displaying them as an artificial graphical user interface (GUI) element. We believe that this enhances the virtual experience, and thus, makes SSVEP-based interaction more applicable. Concretely, these stimuli should blend in a given virtual environment, and thus, are not recognized by the users as stimuli. Therefore, we developed stimuli in form of virtual butterflies with different levels of *shape realism* that elicit SSVEP responses through *flickering* or *flapping* wings. *Flapping* wings can work similar to GSS [7] or spinning icons [43] as they change their angular size while being focused by the user. As prior work hints towards a connection between visual comfort and classification accuracy [7], we believe that especially in VR it is important to find a suitable trade-off between classification accuracy and visual appearance.

# 3 GENERAL APPROACH

In this section, we introduce our approach to well-suited and realistically shaped stimuli for SSVEP-based interaction in VR. We first extracted nine commonly used frequencies from the literature. Out of these nine frequencies, we determined three frequencies with the highest detection accuracy in our first study with twelve participants. We used these three frequencies in our second study to train a classifier (i.e., SVM) to detect SSVEP responses elicited through our stimuli in form of butterflies with different levels of shape realism and different wing animations – with *flickering* or *flapping* wings. Further, we obtained subjective feedback on perceived realism, movement, visual pleasantness, and ability to focus on the stimuli. During our studies, we followed the local ethical process.

*Identification of Suitable Frequencies.* To design our butterfly-shaped stimuli, we first obtained a frequency range that is frequently used for SSVEP responses. Therefore, we conducted a brief literature review. We queried the *ACM Digital Library* on May 25,

**Table 1: Overview on frequencies used in the retrieved literature using the query *"Brain-Computer Interface"* AND *"SSVEP"* on the *ACM Digital Library*. A * indicates increments of 2 Hz and ** indicate increments of 0.2 Hz in a range of frequencies.**

| Stimulus Type | Frequency Range | Stimulus Device | Reference |
|---|---|---|---|
| Flickering Rectangles | 5, 6, 7.5, 8.33, 8.57, 10, 12, 12.5, 15, 6-20 Hz | Display | [20, 21, 23, 47] |
| Flickering Squares | 6.0, 6.25, 6.67, 7.50, 7.57, 8, 8.6, 9, 10 Hz | Display | [30, 39, 44, 45] |
| Flickering Circles | 6, 6.67, 7.5, 8.57, 10, 12, 15 Hz | Display | [9, 13–15] |
| Flickering Checkerboard | 8, 9, 11, 12, 15, 20, 25 Hz | Display | [34, 63] |
| Flickering Square | 6.25, 8, 9, 10 Hz | Smartphone | [40] |
| LED | 5.783, 6.75, 8.65, 11-19, 30-48 Hz* | LEDs | [18, 53, 54] |
| Flickering Squares | 8-15.8 Hz** | VR-HMD | [32] |
| Flickering VR Objects | 42, 43, 44, 45 Hz | VR-HMD | [2] |
| Flickering Rectangles | 4, 6, 8, 9, 10, 11, 12, 13, 15 Hz | AR-HMD | [58, 59] |

2021, with the following search terms *"Brain-Computer Interface" AND "SSVEP"*. We retrieved 92 items from the library. From the results, we selected publications that investigated SSVEP-based interaction and extracted the frequencies used in the experiments. We ignored literature that used BCIs with other interaction paradigms than SSVEP (e.g., P300 [42]). Further, we solely considered *research-articles*s, *short-papers*s, *abstract*s, and *survey*s. We excluded literature that did not report on the frequencies used in their experiments. In Table 1, we present an overview of the gathered frequency ranges together with further details like stimulus type and devices used for stimulus emission.

*Study I: Selection of Suitable Frequencies.* We selected a set of 72 unique frequencies which were used in previous SSVEP studies (see Table 1). Then, we counted the occurrences of all used frequencies and removed all frequencies which occurred fewer than three times. Furthermore, we removed decimal frequencies to avoid interpolation between two frequencies. This resulted in the following nine frequencies: $6Hz$, $8Hz$, $9Hz$, $10Hz$, $11Hz$, $12Hz$, $13Hz$, $14Hz$, $15Hz$. In Study I, our participants viewed squares flickering at each of these frequencies while we measured cortical activity at the occipital lobe. Twelve participants took part in this study. After the study, we compared the classification results of each possible frequency triplet. After extracting the triplet with the highest classification accuracy, we continued with our second study in which we manipulated the appearance and animation of our butterfly-shaped stimuli using this frequency triplet.

*Study II: Evaluation of Accuracy an Appearance.* We developed three SSVEP stimuli in the form of butterflies with increasing realism regarding their shape and stimulus type (see Figure 1). We intended to increase the realism of the shapes by alternating their wings from square wings (i.e., *low shape realism*), similar to the square used in Study I or previous work [24, 49, 51], over round wings (i.e., *moderate shape realism*) to real wing contours (i.e., *high shape realism*) of a real butterfly. The use of a butterfly was inspired by previous research that successfully evaluated navigation using SSVEP butterflies in a non-VR environment [27]. Butterflies have relatively large wings compared to their body. These large wings allowed us to create a large stimulus area to excite the retina of an observer. To elicit SSVEP responses through the natural movement of the butterfly, we animated the wings to move up and down at specific frequencies, and thereby, changing their angular size. This elicits SSVEP response similar to spinning icons [43] or GSS stimuli [7]. Our butterflies flapped their wings from 90° upwards to 90° downward and back at a specific frequency.

## 3.1 Study Apparatus

For both studies, we developed a VR app in *Unity3D*. The VR app placed the user in a dark room with gray walls. Thereby, we could reduce the influence of external factors and focus on the evaluation of the tested stimuli. The VR app was configured in *Unity3D* to display either a square stimulus or the butterflies with different levels of shape realism and wing animation. In Study I, the app displayed a black and white flickering square measuring $0.4m \times 0.4m$ at a $1m$ distance (see Figure 2). The square was colored white with a black frame. The white area covered $1.162m^2$. For reference, $1m$ in VR corresponds to $1m$ in reality. In Study II, the app displayed our butterflies once at a time (see Figure 1). One wing of our *low, moderate, and high shape realism* butterflies encompassed a white area of $1.16m^2$, $0.88m^2$, and $0.60m^2$ respectively. From the center of our butterflies to the farthest point on the wings we measured a distance of $48cm$, $40cm$, and $33.6cm$ for our *low, moderate, and high shape realism* butterflies. To run the VR app, we connected an *Oculus Quest 2* to a PC via an *Oculus Link* cable. This allowed us to operate the VR app from outside the VR-HMD while monitoring the EEG signal (see Figure 2, left) and simultaneously the first-person view of the participants (see Figure 2, right). In our setup, we had a constant refresh rate of $72Hz$. The stimulation signal was modeled using a square wave in a custom shader for best performance. We measured around $7.33lx$ illuminance per eye emitted by the HMD using a photometer[2] when showing a *low shape realism* butterfly when it was rendered fully white and wings were spread to the maximum (see Figure 1). When showing a *moderate shape realism* and *high shape realism* butterfly, we measured $6.33lx$ and $5.67lx$ respectively. As the maximum frequency was $15Hz$ in Study I, we were certain that our stimuli were presented properly as the refresh rate of our VR-HMD was about five times higher ($72Hz$) than $15Hz$. Additionally, we recorded the stimuli with an external highspeed camera[3] from within the VR-HMD. By checking the recording frame by frame, we were certain that our stimuli were presented properly.

We recorded the EEG signal using an *OpenBCI Ganglion EEG board* since its electrodes can be easily integrated into VR headsets

[2]VOLTCRAFT LX-10 Photometer, range: 0 - 199900 lx
[3]Camera Model: ELP-USBFHD08S, max. frame rate: 720p@260 fps

**Figure 2: For our study apparatus, we use an Oculus Quest 2 VR-HMD and OpenBCI Ganglion EEG board. Left: A participant wearing a VR-HMD focused a stimulus. Right: The same participant in VR focusing an a black and white flickering square.**

while maintaining a signal recording quality comparable to medical-grade devices [10]. The *Ganglion* operates with a $200Hz$ sampling frequency and has 4 input channels. We used 2 out of the 4 available channels to sample EEG signals. We placed the electrodes on the occipital lobe of the participant at *POz* and *Oz* according to the 10-20 system [17]. We placed the ground electrode on the right earlobe and the reference electrode on the left earlobe in Study I. However, we changed the position of the reference electrode to *Cz* since it improved the signal quality in Study II. We streamed the EEG signal from the *OpenBCI GUI* to our VR app using the open sound protocol (OSC)[4]. The VR app annotated the EEG signal with the current stimulus frequency and stores it in comma-separated files for later analysis and classifier training.

## 3.2 Data Processing and Machine Learning Approach

In the following, we introduce our data pre-processing and machine learning approach to classify the SSVEP responses elicited through our stimuli.

*3.2.1 Preprocessing and Classifier Training: Study I.* We divided the raw EEG signal into epochs relating to the displayed frequencies. We average the raw EEG signal of the electrodes *POz* and *Oz* to obtain a single signal. This is a known method to denoise biomedical signals [41]. One second of data was removed from the beginning and end of each trial to remove signals that are unrelated to cortical activity. Each epoch was high pass filtered at $0.1Hz$ and low pass filtered at $40Hz$. We have intentionally selected a cutoff at $40Hz$ to include harmonic frequencies. Including additional harmonic frequencies is known to increase the classification accuracy due to the occurrence of more robust features [38]. We performed a *Short-time Fourier transform* on two-second slices with an overlap of one second. The obtained frequency bins per second were labeled with the displayed SSVEP stimulus [6], representing the feature vectors used to train a Support Vector Machine (SVM) [8]. We performed a grid search to find the optimal hyperparameters for the test set [19]. We evaluate the classifier performance through cross-validation with $k = 10$, where $k - 1$ folds were iteratively used for training and the remaining fold was used for evaluation.

*3.2.2 Preprocessing and Classifier Training: Study II.* In Study II, we recorded the EEG data while displaying our butterfly stimuli. We obtained raw EEG recordings along with annotations with the respective frequency of the stimuli. Then, we applied the following pre-processing steps. First, we averaged the signal similar to Study I (*POz* and *Oz*, Ref *Cz*, GND *right earlobe*) and created buckets with 200 samples each by using a sliding window approach with a step size of one. As we repeated the stimuli exposure three times per frequency, we separated the second block from the first and third block to use it for testing. Next, we normalized each bucket using *zero-mean normalization* and applied a band-pass filter from 0.1 to $40Hz$. We then computed the *Fast-Fourier Transform* of each bucket. As a result, we obtained a training set and test set of transformed buckets for each frequency. We then trained our SVMs with the training set and tested it on the test set.

[4]Open Sound Control (OSC). www.opensoundcontrol.org, last checked on March 11, 2022

## 4 STUDY I: SELECTION OF SUITABLE FREQUENCIES

The goal of the first study was to determine a combination of best-performing frequencies identified by previous research. In this study, participants were viewing abstract SSVEP stimuli using the frequencies from the literature while we recorded cortical activity.

## 4.1 Study Design

We conducted a within-subjects laboratory user study in VR using our previously described apparatus to compare the most reported frequencies (see Subsection 3.1). In this study, we showed a square stimulus in the center of the participants' field of view, flickering between black and white with different frequencies. Our only independent variable was *frequency* with the nine levels (*6Hz*, *8Hz*, *9Hz*, *10Hz*, *11Hz*, *12Hz*, *13Hz*, *14Hz*, and *15Hz*). Each frequency was displayed three times to the participants; each for ten seconds with additional prior five seconds in which no stimuli was visible. Participants received a short break after each block. The frequencies displayed in each block were counterbalanced using a Latin square design. During each trial, we measured the participants' cortical activity as described in Subsection 3.1. This results in an overall data collection of 30 seconds per participant and per frequency. In our analysis, we focused on the selection of a triplet yielding the highest detection accuracy. We have selected an overall number of three frequencies since previous research found this number of frequencies suitable for interaction [27]. This study addresses the following research question:

**RQ1**: Which triplet combination of the nine frequencies provide the best detection accuracy for SSVEP in VR?

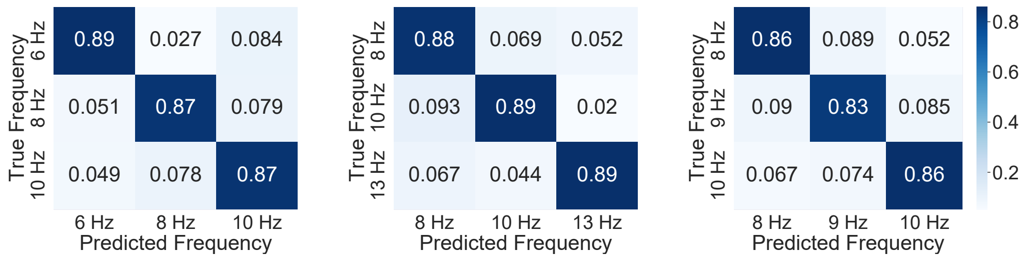


**Figure 3: Confusion matrices of the three frequencies yielding the highest accuracies. Left: Classification using 6, 8, and 10$Hz$. Middle: Classification using 8, 10, and 13$Hz$. Right: Classification using 8, 9, and 10$Hz$.**

## 4.2 Procedure

Participants were introduced to the purpose and procedure of the experiment at the beginning of the study. We asked participants if they are affected by neurological disorders (e.g., epilepsy) to ensure participation without risk. Then, we asked participants to fill out a demographics questionnaire. Next, we started to set up the devices used in our experiment. First, we helped participants to mount the passive gold cup electrodes on their heads (see Subsection 3.1). We conducted several measures to verify the correctness of our data collection. We ensured that the impedance of each electrode is lower than 25$kOhm$ throughout the experiment. We asked participants to close their eyes to verify measuring the visual cortex by observing spikes in the alpha band. Then, we stimulated the participants with three test frequencies (i.e., 7$Hz$, 8$Hz$, and 12$Hz$) through a 240$Hz$ LED display[5]. We chose these frequencies for initial testing as they worked well during the development process. Through the *OpenBCI GUI*, we observed the signal in the frequency domain to assure the correct SSVEP responses. We continued with helping the participants mount the VR-HMD when responses to all three frequencies were visible. We checked the impedance again after mounting the HMD. If the impedance remained the same, we stimulated the participants in VR with our three test frequencies and checked again the responses to the stimuli in the frequency domain. We continued with the data recording described in Subsection 4.1 when the verification steps were accomplished. Each participant took around 30 minutes to finish the study.

## 4.3 Participants

We recruited 12 volunteers (8 male, 4 female, 0 other), aged between 23 and 33 years ($M = 28.3$, $SD = 3.3$). Three participants had corrected-to-normal vision and one participant reported colorblindness. None of the participants stated any neurological disorders and every participant assured us to not be affected by epilepsy. Participants were asked to rate their experience with VR on a 7-point Likert scale (1: no experience; 7: expert-level experience). Most participants stated that they were familiar with VR ($Md = 4.5$, $IQR = 2.5$).

## 4.4 Classification Results

We assessed the classification accuracy for every possible combination of three frequencies that we found in the literature. This resulted in 84 distinct frequency combinations. We applied the classification procedure described in Subsection 3.2 separately on each participant by training and evaluating a Support Vector Machine (SVM). We averaged the resulting performance metrics of each classifier and for each frequency combination to obtain the best performing frequencies. On average, we received 106 feature vectors per frequency combination and per participant for training the SVMs. Each feature vector included 40 features (i.e., one for each frequency power bin). A grid search [19] suggested a radial basis function kernel and a regularization parameter of $C = 2$ for evaluation.

We calculate the precision, recall, and $F_1$ scores as performance metrics for each participant. In addition, we calculate the accuracy as the number of correct predictions divided by the number of total predictions. We multiplied the accuracies by 100 to obtain percentages. We then calculated the accuracy, precision, recall, and $F_1$ scores for each participant and for each frequency triplet. We then averaged the accuracy, precision, recall, and $F_1$ scores over all participants to obtain four single performance metrics. The accuracy ranged between 67% (i.e., the lowest accuracy was achieved by {6, 9, 12}) and 83% (i.e., the highest accuracy was achieved by {6, 8, 10}). Table 2 summarizes the five best performing frequency triplets along with their accuracies. Figure 3 shows the confusion matrices of the three best performing frequency triplets.

**Table 2: The five frequency triplets achieving the highest accuracy.**

| Frequencies (in Hz) | Accuracy | Precision | Recall | $F_1$ Score |
|---|---|---|---|---|
| (6, 8, 10) | 83% | .80 | .80 | .80 |
| (8, 10, 13) | 83% | .80 | .80 | .80 |
| (8, 9, 10) | 82% | .79 | .78 | .78 |
| (8, 10, 12) | 82% | .78 | .78 | .78 |
| (6, 8, 14) | 82% | .81 | .81 | .81 |

We observed the highest classification accuracy for {6, 8, 10} with an accuracy of 83% ($F_1$ = .80), followed by {8, 10, 13} with an accuracy of 82% ($F_1$ = .80), and {8, 9, 10} with an accuracy of 82% ($F_1$ = .78). Overall, high accuracies were achieved for all frequency triplets. An exemplary spectrogram of one participant (P7) shows the distinct pattern of the elicited SSVEP responses in Study I (see Figure 4).

## 4.5 Discussion

We conducted Study I to obtain a set of SSVEP frequencies for reliable classification using machine learning. Therefore, participants

[5] Acer LED Display. https://www.acer.com/ac/en/US/content/predator-series/predatorxb2, last checked March 11, 2022.

focused on black and white flickering squares in VR. We used nine common frequencies selected from related work while we recorded cortical data. Afterward, we examined the classification accuracy by evaluating all possible frequency triplets.

Our results show that all frequency triplets provide sufficient accuracy over the expected chance level, where our reported accuracies are similar to values reported by past research [6]. However, a particular difference in accuracy exists between triplets with low accuracy (e.g., {6, 9, 12} reaching 66%) and triplets with high accuracy (e.g., {6, 8, 10} reaching 83%). We noticed that high-scoring frequency triplets have a large number of harmonics in the feature space (i.e., $0.1Hz - 40Hz$). Previous work stated that classification accuracy improves when non-conflicting harmonics are present in the data set [38]. However, this comes at the cost of reducing the number of usable frequencies. For example, $6Hz$ and $12Hz$ share harmonics multiple times. This explains why triplets such as {6, 9, 12} showed a poor performance. Although the best triplets conflict with harmonics as well with higher frequencies, they still provide enough distinct features resulting in a more robust classification. Reaching high classification accuracies is a major objective of BCI research to maintain a solid interaction experience for the user. Hence, we decided to continue with the {6, 8, 10} frequency triplet.

# 5 STUDY II: EVALUATION OF APPEARANCE AND ACCURACY

In our second study, we continued with the three frequencies which we could classify most accurately in Study I. The goal of the second study was to investigate the influence of factors that introduce more realism to the used stimuli. Inspired by previous work [27], we selected a butterfly to elicit SSVEP responses. We evaluated butterflies with different shapes. Here, we had three levels of *shape realism* and different *stimulus types* which are *flickering* or *flapping* wings.

## 5.1 Study Design

To investigate the influence of different levels of shape realism on SSVEP stimuli, we conducted a within-subjects laboratory user study in VR using our described apparatus (see Subsection 3.1). We used a repeated-measures design to examine the influence of two independent variables on the accuracy of the visually evoked potential. Our independent variables were *shape realism* with three levels (*low* vs. *moderate* vs. *high*) and *stimulus type* with two levels (*flicker* vs. *flap*), resulting in overall six conditions. In each condition, we tested three different frequencies ($6Hz$ vs. $8Hz$ vs. $10Hz$). The conditions were counterbalanced using a Latin-square design. Within each condition, we recorded three trials of each *frequency*. During each trial, we measured participants' EEG response as described in Subsection 3.1.

We posed the following research question for Study II:

**RQ2**: Which combination of *shape realism* and *stimulus type* results in the highest classification accuracy and which one is perceived as most realistic?

Given the purpose of our study, we posed the following hypotheses:

$H_1$ : We expect that the *stimulus type flicker* leads to better classification accuracy than *flap* because the state change for *flicker* is binary while *flap* is a continuous motion that contains mostly intermediate states while the extremes are visible for only a short time.

$H_2$ : We hypothesize that a higher *shape realism* results in participants perceiving the stimulus as more realistic.

## 5.2 Procedure and Participants

In the beginning, we introduced participants to the purpose and procedure of our second study. Thereafter, we asked participants for their consent to the study conditions. We started the study by setting up the devices involved (see Subsection 3.1). Then, we ensured that the impedance of each electrode was lower than $25kOhm$ and assured that we obtain a clear EEG signal similar to the procedure in Subsection 4.2. If everything looked as expected, we continued with the main part of the study.

For each participant, we tested the six conditions with different levels of *shape realism* and *stimulus type* in our VR app to obtain training data and evaluate the classification performance (see Figure 1). We configured our VR app to display the butterfly stimuli of each condition in blocks of 10 seconds for each of the three selected frequencies from our first study. Before each block, there was a break with a duration of 5 seconds. This was repeated three times to obtain a 30 seconds recording of EEG data for each frequency per stimulus. We stored the raw EEG signal in comma-separated files including annotations of the displayed frequency. After each trial, we asked participants to rate 7-point Likert statements to assess their subjective perception of the stimulus and gathered informal feedback of the participants in semi-structured brief interviews. Overall, the study took 45 minutes on average. We recruited the same 12 participants that also participated in Study I (see Subsection 4.3).

## 5.3 Results

For descriptive statistics, we report mean (*M*), median (*Md*), and interquartile-range (*IQR*). Effect sizes of performed statistic tests are reported with r (r=0.1 small effect, r=0.3 medium effect, and r=0.5 large effect).

*5.3.1 Effects on Classification Accuracy.* For classification accuracy (in percentage), we report the F1 scores of the classifiers as the harmonic means of recall and precision. We observed that recall and precision performed similarly across conditions, and thus, decided that the F1 score is a good measure to reflect on both. In the following, we compare the F1 scores for the trained classifiers, to understand how they are affected by our independent variables. For each participant, we trained one classifier for each condition (*shape realism* x *stimulus type*) and took their F1 scores for our analysis. We adjusted the p-values with a Bonferroni correction considering all comparisons. For adequate statistical power, we investigated our independent variables separately (i.e., fewer overall comparisons result in less *p*-adjustment).

*Shape Realism.* We consider the effect of *shape realism* on the F1 score for each level of *stimulus type* individually (see Figure 5). For *flicker*, the median (interquartile-range) F1 scores for the levels of *shape realism* are (in desc. order): *flicker+low*=78.7% (IQR=17.6%), *flicker+moderate*=71.6% (IQR=31.4%), and *flicker+high*=67.5% (IQR=29.3%). A Friedman test revealed a significant effect ($\chi^2(2)$=7.39,

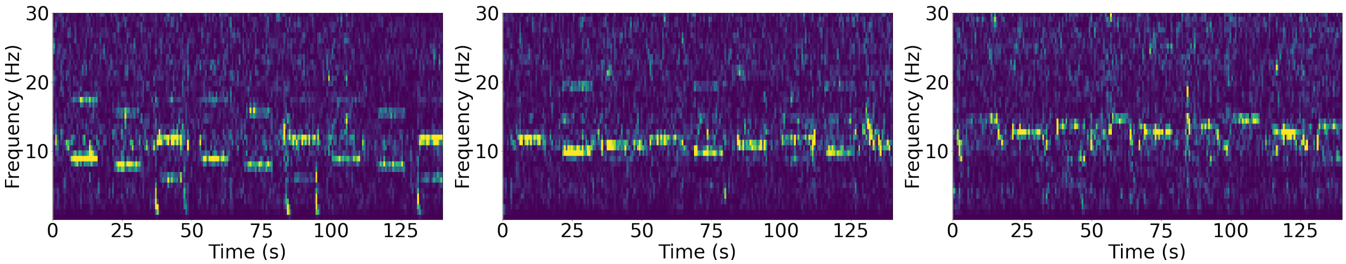


**Figure 4: Spectrogram of P7 in Study I. Left: Low frequencies (8, 9, 10$Hz$). Middle: medium frequencies from (10, 11, 12$Hz$). Right: High frequencies (13, 14, 15$Hz$). When the participant was exposed to a 10$s$ stimulus one can see the higher amplitude of the stimulus frequencies and their harmonics. In between, when now stimulus was applied in the 5$s$ break, gaps are visible.**

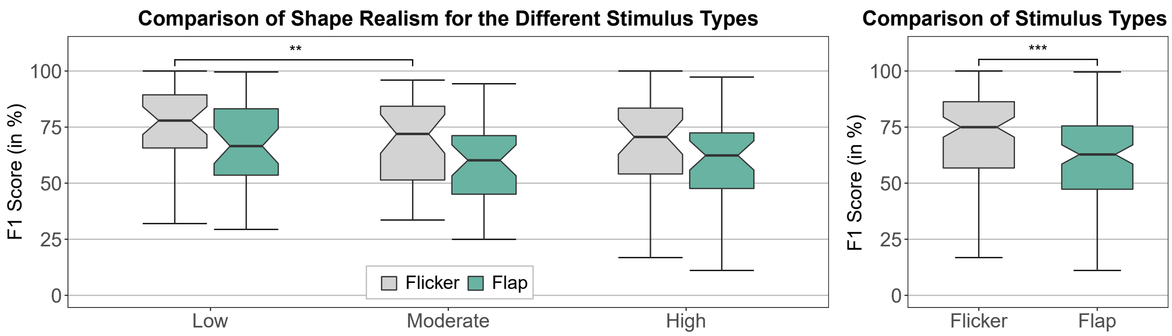


**Figure 5: Comparison of recall rates. Left: comparison of the different levels of shape realism for each of the two stimulus types. Right: comparison between the investigated stimulus types: *flicker* and *flap*. The significance levels are: *(<0.05), **(<0.01), and ***(<0.001).**

p=0.025, N=12). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed a significant difference between *flicker+low* and *flicker+moderate* (W=539, Z=3.24, p=0.005, r=0.38), meaning *low* works significantly better than *moderate* for *flicker*. For *flap*, the median (interquartile-range) recall rates for the levels of *shape realism* are (in desc. order): *flap+low*=70.0% (IQR=29.4%), *flap+high*=60.1% (IQR=21.0%), and *flap+moderate*=57.7% (IQR=17.8%). A Friedman test did not reveal a significant effect ($\chi^2(2)$=1.06, p=0.590, N=12).

*Stimulus Type.* We consider the effect of *stimulus type* on the F1 score. The median (interquartile-range) recall rates for the different levels of *stimulus types* are (in descending order): *flicker*=72.1% (IQR=27.4%) and *flap*=62.0% (IQR=22.8%) (see Figure 5). As we do not assume normality and compare two matched groups within-subjects, we directly performed a Wilcoxon Signed-rank test. Here we found a significant effect of *stimulus type* on recall rate (W=4229, Z=3.94, p<0.001, r=0.27). This indicates that *flicker* has significantly better performance than *flap*.

*5.3.2 Subjective Ratings.* After each condition (*shape realism* x *stimulus*), we asked participants to rate two statements with 7-point Likert items (1=strongly disagree, 7=strongly agree). All ratings are shown in Figure 6.

*Shape Realism.* For the first statement, "the stimulus shape looked realistic," the ratings for the conditions are reported in Figure 6.

**Table 3: Pairwise comparisons of conditions concerning subjective responses concerning *shape realism*.**

| Comparison | | | W | Z | p | r |
|---|---|---|---|---|---|---|
| *flicker+low* | vs. | *flicker+moderate* | 6 | -2.43 | 0.047 | 0.50 |
| *flicker+low* | vs. | *flicker+high* | 1.5 | -2.96 | 0.004 | 0.60 |
| *flicker+moderate* | vs. | *flicker+high* | 2.5 | -2.76 | 0.015 | 0.56 |
| *flap+low* | vs. | *flap+moderate* | 5 | -2.26 | 0.076 | 0.46 |
| *flap+low* | vs. | *flap+high* | 0 | -3.09 | 0.001 | 0.63 |
| *flap+moderate* | vs. | *flap+high* | 0 | -2.98 | 0.006 | 0.61 |

A Friedman test revealed a significant effect of condition on rating ($\chi^2(5)$=35.49, p<0.001, N=12). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences (see Table 3).

*Natural Movement.* For the second statement, "the stimulus movement looked natural," the ratings for the conditions are reported in Figure 6. Grouped by the *stimulus type*, the median (interquartile-range) ratings are (in descending order): *flicker*=2 (IQR=3) and *flap*=3 (IQR=3). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed a significant difference between *flicker* and *flap* (W=129.5, Z=−2.07, p=0.037, r=0.24), meaning the movement of *flap* looked significantly more realistic than *flicker*.

*5.3.3 Participants' Feedback.* Besides the statement ratings on realism, movement, we gathered informal feedback from our participants after each trial. We used thematic analysis to group the feedback of the participants. Two researchers coded statements independently resulting in 82 open codes. Next, we employed an affinity diagram [16] of the open codes and organized the codes into groups, which were then further refined into themes using an online whiteboard[6].

*General Feedback.* Participants described the butterfly with *low* or *moderate* shape realism as glaring, overwhelming, or exhausting: *"This was exhausting [...] it just blinked and did not move. I stared into whiteness."* (P9 on moderate shape realism, flickering butterfly). Regarding the shape of the butterfly, participants stated that the two wings made were distracting and made it difficult to focus on the stimulus: *"It looked like a TV. And two different sides to focus on are difficult because you don't know where to focus. When I focused on one side the other side was distracting me"* (P5, on low shape realism, flickering butterfly). Also, the participants missed textures or colors on the wings of the butterflies, *"Colorful wings would be cool! Then I wouldn't even recognize this as a stimulus."* (P8 on high shape realism, flapping butterfly) and P6 commented on a high shape realism, flapping butterfly: *"This was the best stimulus so far. I had no problems focusing on the stimulus and the movement felt like the natural movement of a butterfly."*

*Flickering.* On the one hand, our participants stated that flickering was easier and less chaotic to focus on: *"This time I looked more at the wings, it was easier to focus on [the flickering wings] than the moving wings[...]."* (P1, on high shape realism, flickering butterfly). On the other hand, they stated that flickering wings are not realistic: *"Flickering is easier to focus but not that realistic"* (P8 on a moderate shape realism, flickering butterfly).

*Flapping.* On the one hand, the participants stated that the wings flapped too fast: *"It looked more like a flicker book to me."* (P3 on a low shape realism, flapping butterfly). Participants used the butterflies' body as a reference point to focus when they were flapping their wings: *"[...] When the wings were in motion the body was easier to focus."* (P1, on high shape realism, flickering butterfly). Some participants perceived the flapping wings as less realistic and stated that the wing motion negatively influenced the focus on the butterfly. On the other hand, the participants liked the flapping wings and stated that the stimulus was less intense than flickering wings: *"The flickering, especially at high frequencies, was unpleasant. This was not the case when the wings were flapping."* (P3 on a low shape realism, flickering butterfly) and *"The realistic, flapping [butterfly] was the most pleasant. Through the flap, it is more pleasant in general and not as intense [as flickering]."* (P2, on high shape realism, flapping butterfly). One participant stated that through a more realistic shape the flapping motion appeared more realistic: *"[...] the shape of the wings appeared more realistic when they were in motion [...]."* (P1 on moderate shape realism, flapping butterfly).

*Focus.* Several participants stated that the butterfly's body helped them keep the focus on the stimuli while the wings were flapping. They stated that the fixed body was easier to focus on than the moving parts: *"It was easier to focus! I could concentrate on the [butterfly's] body and the stimulus was around it."* (P2, on moderate shape realism, flapping butterfly). Also, the form of the wings helped the participants to maintain focus when the butterfly's wings were flickering: *"The round form makes it easier to focus a circle [...]."* (P1 on moderate shape realism, flickering butterfly). The realistic contours were perceived similarly: *"Through its contours, the stimulus was better to focus."* (P1, on high shape realism, flickering butterfly). In contrast to this, participants stated to have problems focusing on the butterflies because they were more realistic: *"The stimulus resembles the butterfly flying and I focused on the shape and details of the butterfly. I did not really know where to focus the stimulus."* (P8, on moderate shape realism, flapping butterfly). Several participants stated that the butterflies with realistic contours were distracting them because they investigated the wings: *"Through the complex form, I tended to investigate it [...]. It was difficult to keep the focus centered on the butterfly because I felt urged to scan the butterfly with my eyes."* (P1, high shape realism, flickering butterfly).

[6]Miro. https://miro.com, last retrieved March 11, 2022

*Levels of Shape Realism.* The shape realism was perceived differently by our participants. The butterfly with a low shape realism was perceived as *"machine-like"* (P9 on a low shape realism, flapping butterfly), it resembled an *"[...] old TV with an antenna."* (P1 on a low shape realism, flickering butterfly) or it would fit in video games like *"Minecraft"* (P8 on a low shape realism, flickering butterfly). Unlike most participants, one appreciated the butterfly with round wings the most: *"I think that the round wings looked more realistic than the wings with real contours. I find them more pleasant"* (P9). Our participants stated several times that they liked a high level of shape realism: *"The shape of the wings looks realistic here. But I missed flapping wings. This would appear more realistic to me."* (P8, on high shape realism, flickering butterfly). After perceiving the high shape realism butterfly with flapping wings, P8 added *"I think this looks like a butterfly! The shape matches and it's flapping."*

## 5.4 Discussion

We assessed the different levels of *shape realism* and *stimulus type* of different butterfly configurations in terms of classification accuracy and subjective appearance in Study II. In the following, we discuss the results.

*Classification Accuracy.* We found that *flickering* wings resulted in a significantly higher F1 score. This shows that *flickering* outperforms *flapping* wings in terms of classification accuracy. Therefore, we can accept our hypothesis $H_1$. We argue that while wings are flapping, the participants could see the white area of the wings growing and shrinking similar to GSS [7] or spinning stimuli [43]. This results in a less effective stimulus than black and white flickering wings and is in line with previous observations [7].

When we analyzed the influence of the shape on the classification accuracy, we observed a higher median for a *flapping* butterfly with *high* shape realism than for a butterfly with *moderate* shape realism even though it comprises a smaller white wing area. There is a tendency towards the claims from the literature that suggest that stimuli preferred by users perform better in terms of classification accuracy [7]. But we could not show an effect here. Still, a *flapping low shape realism* butterfly performed best. Here, we must acknowledge that it had the largest wing area. For *flickering* wings, we observed different results. Here, a *flickering* stimulus

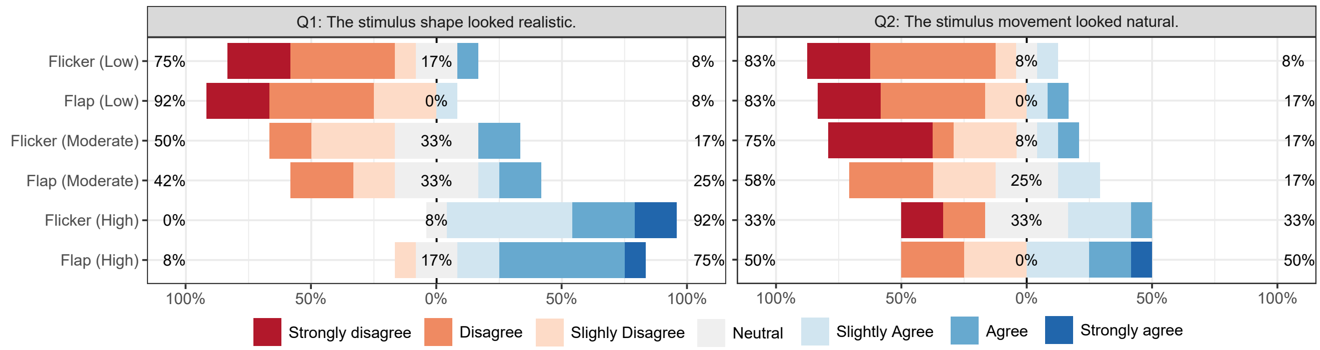


**Figure 6: Subjective ratings with 7-point Likert-items for each condition tested. Left: subjective ratings for the shape realism of the butterfly shape. Right: subjective ratings for the naturalness of the butterfly movement.**

with *low* shape realism performed significantly better than a *moderate* butterfly. This could be attributed to the smaller white area of the round wings. Further, our participants reported that they scanned interestingly shaped wings and therefore had some difficulty maintaining focus. Participants reported that focusing on the body of the butterfly with round wings was easier due to the gap between its body and its wings. Therefore, the stimulating wings might not be in full focus of the participants. This could result in less classification accuracy. Focusing on details of *flapping* wings might be difficult for users as they move quickly. This could be a benefit of flapping stimuli as users do not tend to focus on details, instead, they look at the entire stimulus. For *flapping wings*, we did not find any significant differences. We conclude that when designing stimuli it is important to consider how users perceive the difficulty to keep their focus on the stimuli. Especially during long-term interaction, this could have a negative impact.

*Shape Realism.* In terms of appearance, our participants rated the *flickering* butterflies with realistic contours to appear significantly more realistic than the butterflies with *low* or *moderate* shape realism. We observed the same for butterflies with *flapping* wings. Therefore, we accept $H_2$.

*Flickering vs. Flapping.* Subjectively, our participants perceived the flapping butterflies as significantly more natural than butterflies with flickering wings. However, flickering stimuli resulted in higher classification accuracy. Here, we face a trade-off in terms of classification reliability and stimulus appearance. In VR, realism is an important factor for users to immerse in virtual worlds. Therefore, we argue that in some cases it is acceptable to employ stimuli that appear more realistic, in our case *flapping* wings, and at the same time sacrifice a certain percentage of classification accuracy. For example, in a game an animal that elicits SSVEP responses could look at the player but if the detection fails the player would not notice a big difference but overall the game would be more interactive. Also, long-term interaction might benefit from visual pleasant stimuli to mitigate adverse side effects on the user like mental load [64] or visual fatigue [36, 37].

When we compared the classification accuracy of our stimuli that were preferred by our participants in terms of *shape realism* and *stimulus type*, we did not observe mixed classification accuracy of stimuli that were rated more visually pleasing than others. We could not show that a visually more pleasant stimulus positively influences the classification accuracy of SSVEP stimuli as suggested by the literature [7] across all levels of shape realism. Still, a *high* shape realism butterfly achieved a slightly higher classification accuracy than a *moderate* one even with a smaller stimulus area that excites the users' retinas. Here, we suggest further investigating factors of our butterfly stimuli such as the shape, colors as well as textures, or wing speed. Overall, our stimuli that appeared more pleasant to our participants still achieved satisfactory classification accuracies applicable in a wide array of apps.

## 6 GENERAL DISCUSSION

In the following, we discuss the results from our studies along with limitations and future research suggestions.

*Trade-off Between Detection Accuracy vs. Stimuli Appearance.* Our findings suggest that there is a trade-off between the performance and the appearance of our stimuli. This trade-off should be considered when integrating such stimuli in VR. If the interaction must be robust in terms of classification accuracy, a *flickering* stimulus might be well-suited. When appearance or proper CGI is more important than performance, stimuli with matching animation, in our case *flapping* wings, could be integrated into VR as they blend into the surrounding environment through a plausible animation. This would help to preserve the narrative of the virtual experience. We obtained promising feedback from our participants, who stated that more realism might further disguise the fact that our butterflies are SSVEP-eliciting stimuli. We conclude that our findings are transferable findings to other real or fictional animals with wings including but not limited to birds, flies, bees, or dragons. Beyond these, it is not clear how generalizable our findings are to other objects. However, a wide array of research demonstrates the potential of SSVEP for non-flickering stimuli [2, 43], including ours, suggesting great potential for blending SSVEP stimuli in VR in general.

*Level of Detail and Perception.* Not all VR apps rely on complex graphic pipelines, photo-realistic details, or a high level of detail. Therefore, we argue that stimuli with a less detailed shape, such as the butterfly with round wings, could serve as SSVEP stimuli in VR, which are built of low poly meshes with a minimal number of details. Overall, our participants rated the butterfly with real wing contours as most realistic but also stated that butterflies with round wings could have use cases. Here, one interesting statement regarding the butterfly with square wings caught our attention. For games such as

*Minecraft*, a butterfly would blend into the environment well. So, we argue that the expectation of VR users can influence the perception of the SSVEP stimuli if a plausibility illusion is created [46]. This could be further investigated by embedding different stimuli designs into VR with varying graphical properties (e.g., colors and textures). Hence, our findings are valuable in terms of the evaluated levels of shape realism. We believe that it is important to not entirely focus research on the most realistic SSVEP stimuli only but rather on a variety of different levels of each dimension that the SSVEP design space has to offer. There are many parameters that can be adjusted when using such SSVEP stimuli in VR. As we did not investigate the whole design space in this paper, we suggest that our work is an initial starting point that can serve as a basis for future research. Here, the whole spectrum of shapes, different animals or other non-living objects, textures, and colors as well as more types of animations span a humongous design space that needs further investigation.

## 6.1 Limitations

We acknowledge various limitations that could have affected our evaluation. First, our stimuli shaped as butterflies are not applicable in every available VR scenario. For example, users would not expect butterflies in an urban setting. This is a limitation to our specific design. To overcome this, creative ideas of VR developers and designers are needed. We outline some inspirational ideas in the future work section. Further, the detection accuracies and the subjective perception of the stimuli could be influenced by the surrounding VR environment. We placed the users in a dark room. This reduced the influence on external factors and we could focus on the evaluation of the presented stimuli. We argue that the detection accuracies could differ when our stimuli are deployed in more complex VR settings. For example, user movement and different light conditions can influence the interaction with our stimuli. Further, the subjective perception of the stimuli could change. Future research could investigate influencing factors when deploying such stimuli in an end-user VR app. Next, we did not rely on VR-HMDs with integrated eye-tracking. Therefore, we cannot quantify which parts of the butterflies were mostly focused on by the participants and which were not. Using eye-tracking could reveal which areas of our stimuli were focused on the most by our participants. This would help to better understand which features of the stimuli attract the attention of VR users.

## 6.2 Future Work

We showed that *flapping* wings could effectively be used as an SSVEP stimulus in VR. This motivates us to outline promising research objectives regarding SSVEP in VR.

The size of our butterflies was larger than butterflies in reality. This was necessary to ensure that SSVEP responses are large enough to be measured by our EEG device. Future work could investigate to what degree the stimuli can be reduced in size while SSVEP responses can be still measured. The decline of the detection accuracy could be attributed to the decreasing wing size of our butterflies when increasing shape realism or the perception of the participants as suggested by the literature [7]. Future evaluations could take this and other variables into account like stimuli distance or moving stimuli as butterflies tend to fly in jagged trajectories. Shrinking the butterflies to a size users are familiar with can further enhance virtual experiences. Also, swarms of butterflies could be investigated.

With our stimuli, we plan to conduct a study that investigates them in a realistic VR environment. We plan to develop a VR game that uses our stimuli to let the player engage in a playful activity. This allows us to evaluate our stimuli in a real-world scenario. As our participants wished for colorful butterflies, we would introduce colored butterflies and repeat our main study. In the future, different wing patterns will be investigated, similar to pattern reversal stimuli [66]. Our participants stated that the wing motion was sometimes too fast. Here, we could slow down the wings and use a combination of *flickering* and *flapping* wings to elicit SSVEP responses while maintaining realistic wing movement.

SSVEP stimuli that allow triggering events or determine the user's focus in future VR games could be generated through the environment itself. One could consider a car driving through a forest. When the sun is low, the light goes through the forest and is blocked by the trees. Depending on the car's velocity, the light is visible only for a specific moment, resulting in a flickering stimulus. This can be used to trigger events when users focus on a specific area of the environment. An equivalent for room-scale approaches could be a lamp behind a fan. The angular velocity of the fan together with the fan's wings that block the light from the lamp creates a flickering SSVEP stimulus. VR developers could use such mechanisms to ensure that the user focuses on objects of interest.

# 7 CONCLUSION

In this paper, we investigated SSVEP stimuli in VR in the form of butterflies with three levels of realism. To elicit SSVEP responses, we developed two stimuli types: *flickering* and *flapping* wings. To assess their suitability for interaction in VR, we first extracted three suitable frequencies through a brief literature survey and subsequent prestudy. We conducted our main study with the three best performing frequencies to obtain training data to train classifiers and to assess the subjective realism of our stimuli. We showed that our stimuli design in the form of a realistic butterfly with *flapping* wings can be used for SSVEP-based interaction in VR, but is still outperformed by a flickering stimulus. Hence, we argue that the stimuli should be selected based on the VR scenario. If performance is required, stimuli with lower levels of realism should be employed. If stimuli should fit the VR environment and robustness can be sacrificed, then higher levels of realism can be used to enhance the VR experience.

# ACKNOWLEDGMENTS

This work is funded by the German Federal Ministry of Education and Research (01IS21068B).

# REFERENCES

[1] Brendan Allison, Thorsten Luth, Diana Valbuena, Amir Teymourian, Ivan Volosyak, and Axel Graser. 2010. BCI Demographics: How Many (and What Kinds of) People Can Use an SSVEP BCI? *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18, 2 (2010), 107–116. https://doi.org/10.1109/TNSRE.2009.2039495

[2] Alexandre Armengol-Urpi and Sanjay E. Sarma. 2018. Sublime: A Hands-Free Virtual Reality Menu Navigation System Using a High-Frequency SSVEP-Based Brain-Computer Interface. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 1, 8 pages. https://doi.org/10.1145/3281505.3281514

[3] Jonas Auda, Roman Heger, Thomas Kosch, Uwe Gruenefeld, and Stefan Schneegass. 2020. EasyEG: A 3D-Printable Brain-Computer Interface. In *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 70–72. https://doi.org/10.1145/3379350.3416189

[4] Vitoantonio Bevilacqua, Giacomo Tattoli, Domenico Buongiorno, Claudio Loconsole, Daniele Leonardis, Michele Barsotti, Antonio Frisoli, and Massimo Bergamasco. 2014. A novel BCI-SSVEP based approach for control of walking in Virtual Environment using a Convolutional Neural Network. In *2014 International Joint Conference on Neural Networks (IJCNN)*. 4121–4128. https://doi.org/10.1109/IJCNN.2014.6889955

[5] Luzheng Bi, Xin-an Fan, Ke Jie, Teng Teng, Hongsheng Ding, and Yili Liu. 2014. Using a Head-up Display-Based Steady-State Visually Evoked Potential Brain–Computer Interface to Control a Simulated Vehicle. *IEEE Transactions on Intelligent Transportation Systems* 15, 3 (2014), 959–966. https://doi.org/10.1109/TITS.2013.2291402

[6] Sarah N Carvalho, Thiago BS Costa, Luisa FS Uribe, Diogo C Soriano, Glauco FG Yared, Luis C Coradine, and Romis Attux. 2015. Comparative analysis of strategies for feature extraction and classification in SSVEP BCIs. *Biomedical Signal Processing and Control* 21 (2015), 34–42.

[7] Kang-min Choi, Seonghun Park, and Chang-Hwan Im. 2019. Comparison of visual stimuli for steady-state visual evoked potential-based brain-computer interfaces in virtual reality environment in terms of classification accuracy and visual comfort. *Computational intelligence and neuroscience* 2019 (2019). https://doi.org/10.1155/2019/9680697

[8] Corinna Cortes and Vladimir Vapnik. 1995. Support-vector networks. *Machine learning* 20, 3 (1995), 273–297.

[9] Andéol Evain, Ferran Argelaguet, Nicolas Roussel, Géry Casiez, and Anatole Lécuyer. 2017. *Can I Think of Something Else When Using a BCI? Cognitive Demand of an SSVEP-Based BCI*. Association for Computing Machinery, New York, NY, USA, 5120–5125. https://doi.org/10.1145/3025453.3026037

[10] Jérémy Frey. 2016. Comparison of a consumer grade EEG amplifier with medical grade equipment in BCI applications. In *International BCI meeting*.

[11] Ola Friman, Ivan Volosyak, and Axel Graser. 2007. Multiple Channel Detection of Steady-State Visual Evoked Potentials for Brain-Computer Interfaces. *IEEE Transactions on Biomedical Engineering* 54, 4 (2007), 742–750. https://doi.org/10.1109/TBME.2006.889160

[12] Xiaorong Gao, Dingfeng Xu, Ming Cheng, and Shangkai Gao. 2003. A BCI-based environmental controller for the motion-disabled. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 11, 2 (2003), 137–140. https://doi.org/10.1109/TNSRE.2003.814449

[13] Hayrettin Gürkök, Gido Hakvoort, and Mannes Poel. 2011. Modality Switching and Performance in a Thought and Speech Controlled Computer Game. In *Proceedings of the 13th International Conference on Multimodal Interfaces* (Alicante, Spain) *(ICMI '11)*. Association for Computing Machinery, New York, NY, USA, 41–48. https://doi.org/10.1145/2070481.2070491

[14] Hayrettin Gürkök, Gido Hakvoort, Mannes Poel, and Anton Nijholt. 2011. User Expectations and Experiences of a Speech and Thought Controlled Computer Game. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology* (Lisbon, Portugal) *(ACE '11)*. Association for Computing Machinery, New York, NY, USA, Article 53, 6 pages. https://doi.org/10.1145/2071423.2071490

[15] Hayrettin Gürkök, Gido Hakvoort, Mannes Poel, and Anton Nijholt. 2017. Meeting the Expectations from Brain-Computer Interfaces. *Comput. Entertain.* 15, 3, Article 5 (April 2017), 10 pages. https://doi.org/10.1145/2633431

[16] Gunnar Harboe and Elaine M. Huang. 2015. Real-World Affinity Diagramming Practices: Bridging the Paper-Digital Gap. In *Proc. 33rd Annual ACM Conf. Human Factors in Computing Systems*. ACM, 95–104. https://doi.org/10.1145/2702123.2702561

[17] Richard W Homan, John Herman, and Phillip Purdy. 1987. Cerebral location of international 10–20 system electrode placement. *Electroencephalography and Clinical Neurophysiology* 66, 4 (1987), 376–382. https://doi.org/10.1016/0013-4694(87)90206-9

[18] Deanna Hood, Damian Joseph, Andry Rakotonirainy, Sridha Sridharan, and Clinton Fookes. 2012. Use of Brain Computer Interface to Drive: Preliminary Results. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Portsmouth, New Hampshire) *(AutomotiveUI '12)*. Association for Computing Machinery, New York, NY, USA, 103–106. https://doi.org/10.1145/2390256.2390272

[19] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. 2003. A practical guide to support vector classification.

[20] Gan Huang, Jianjun Meng, Dingguo Zhang, and Xiangyang Zhu. 2011. Window Function for EEG Power Density Estimation and Its Application in SSVEP Based BCIs. In *Intelligent Robotics and Applications*, Sabina Jeschke, Honghai Liu, and Daniel Schilberg (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 135–144.

[21] Shihong Huang and Pedro Miranda. 2015. Incorporating Human Intention into Self-Adaptive Systems. In *Proceedings of the 37th International Conference on Software Engineering - Volume 2* (Florence, Italy) *(ICSE '15)*. IEEE Press, 571–574.

[22] Han-Jeong Hwang, Jeong-Hwan Lim, Young-Jin Jung, Han Choi, Sang Woo Lee, and Chang-Hwan Im. 2012. Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard. *Journal of Neuroscience Methods* 208, 1 (2012), 59–65. https://doi.org/10.1016/j.jneumeth.2012.04.011

[23] Christoph Kapeller, Christoph Hintermüller, and Christoph Guger. 2012. Usability of Video-Overlaying SSVEP Based BCIs. In *Proceedings of the 3rd Augmented Human International Conference* (Megève, France) *(AH '12)*. Association for Computing Machinery, New York, NY, USA, Article 26, 3 pages. https://doi.org/10.1145/2160125.2160151

[24] Bonkon Koo, Hwan-Gon Lee, Yunjun Nam, and Seungjin Choi. 2015. Immersive BCI with SSVEP in VR head-mounted display. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 1103–1106. https://doi.org/10.1109/EMBC.2015.7318558

[25] Thomas Kosch, Albrecht Schmidt, Simon Thanheiser, and Lewis L. Chuang. 2020. One does not Simply RSVP: Mental Workload to Select Speed Reading Parameters using Electroencephalography. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. ACM, New York, NY, USA. https://doi.org/10.1145/3313831.3376766

[26] Rafał Kuś, Anna Duszyk, Piotr Milanowski, Maciej Łabęcki, Maria Bierzyńska, Zofia Radzikowska, Magdalena Michalska, Jarosław Żygierewicz, Piotr Suffczyński, and Piotr Jerzy Durka. 2013. On the Quantification of SSVEP Frequency Responses in Human EEG in Realistic BCI Conditions. *PLOS ONE* 8, 10 (10 2013), 1–9. https://doi.org/10.1371/journal.pone.0077536

[27] Jozef Legény, Raquel Viciana Abad, and Anatole Lécuyer. 2011. Navigating in Virtual Worlds Using a Self-Paced SSVEP-Based Brain–Computer Interface with Integrated Stimulation and Real-Time Feedback. *Presence* 20, 6 (2011), 529–544. https://doi.org/10.1162/PRES_a_00075

[28] Minglun Li, Dianning He, Chen Li, and Shouliang Qi. 2021. Brain–Computer Interface Speller Based on Steady-State Visual Evoked Potential: A Review Focusing on the Stimulus Paradigm and Performance. *Brain Sciences* 11, 4 (2021). https://doi.org/10.3390/brainsci11040450

[29] Yuanqing Li, Jiahui Pan, Fei Wang, and Zhuliang Yu. 2013. A Hybrid BCI System Combining P300 and SSVEP and Its Application to Wheelchair Control. *IEEE Transactions on Biomedical Engineering* 60, 11 (2013), 3156–3166. https://doi.org/10.1109/TBME.2013.2270283

[30] Chang Liu, Minghui Shi, Min Jiang, Meihong Wu, Tingting Chen, Jianting Liu, Enze Wang, Yuan Shi, and Zilong Li. 2020. Approach to Recognizing SSVEP by Combining CCA with Artificial Neural Network. In *Proceedings of the 2020 the 4th International Conference on Innovation in Artificial Intelligence* (Xiamen, China) *(ICIAI 2020)*. Association for Computing Machinery, New York, NY, USA, 15–19. https://doi.org/10.1145/3390557.3394290

[31] Anatole Lécuyer, Fabien Lotte, Richard B. Reilly, Robert Leeb, Michitaka Hirose, and Mel Slater. 2008. Brain-Computer Interfaces, Virtual Reality, and Videogames. *Computer* 41, 10 (2008), 66–72. https://doi.org/10.1109/MC.2008.410

[32] Xinyao Ma, Zhaolin Yao, Yijun Wang, Weihua Pei, and Hongda Chen. 2018. Combining Brain-Computer Interface and Eye Tracking for High-Speed Text Entry in Virtual Reality. In *23rd International Conference on Intelligent User Interfaces* (Tokyo, Japan) *(IUI '18)*. Association for Computing Machinery, New York, NY, USA, 263–267. https://doi.org/10.1145/3172944.3172988

[33] Krupal Sureshbai Mistry, Pablo Pelayo, Divya Geethakumari Anil, and Kiran George. 2018. An SSVEP based brain computer interface system to control electric wheelchairs. In *2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. 1–6. https://doi.org/10.1109/I2MTC.2018.8409632

[34] Iveth Moreno, Emmanuel Batista, Sofía Serracín, and José Serracín. 2019. SSVEP Offline Analysis Procedures for Low Cost BCI Systems. In *Proceedings of the 2019 3rd International Conference on Computer Science and Artificial Intelligence* (Normal, IL, USA) *(CSAI2019)*. Association for Computing Machinery, New York, NY, USA, 59–63. https://doi.org/10.1145/3374587.3374636

[35] Gernot R Muller-Putz and Gert Pfurtscheller. 2007. Control of an electrical prosthesis with an SSVEP-based BCI. *IEEE Transactions on biomedical engineering* 55, 1 (2007), 361–364.

[36] Katsuyuki Murata, Shunichi Araki, Norito Kawakami, Yuko Saito, and Eiko Hino. 1991. Central nervous system effects and visual fatigue in VDT workers. *International archives of occupational and environmental health* 63, 2 (1991), 109–113. https://doi.org/10.1007/BF00379073

[37] Katsuyuki MURATA, Shunichi ARAKI, Kazuhito YOKOYAMA, Kohei YAMASHITA, Toshihiro OKUMATSU, and Shinichi SAKOU. 1996. Accumulation of VDT work-related visual fatigue assessed by visual evoked potential, near point distance and critical flicker fusion. *Industrial health* 34, 2 (1996), 61–69. https://doi.org/10.2486/indhealth.34.61

[38] Gernot R Müller-Putz, Reinhold Scherer, Christian Brauneis, and Gert Pfurtscheller. 2005. Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. *Journal of Neural Engineering* 2, 4 (oct 2005), 123–130. https://doi.org/10.1088/1741-2560/2/4/008

[39] Viswam Nathan, Jian Wu, Chengzhi Zong, Yuan Zou, Omid Dehzangi, Mary Reagor, and Roozbeh Jafari. 2013. A 16-Channel Bluetooth Enabled Wearable EEG Platform with Dry-Contact Electrodes for Brain Computer Interface. In *Proceedings of the 4th Conference on Wireless Health* (Baltimore, Maryland) *(WH '13)*. Association for Computing Machinery, New York, NY, USA, Article 17, 2 pages. https://doi.org/10.1145/2534088.2534098

[40] Sandra Ebele Nwachukwu, Minghui Shi, Chang Liu, Xiangqian Liu, Changle Zhou, Fei Chao, Min Jiang, Penglin Kang, and Zilong Li. 2019. An SSVEP Recognition Method by Combining Individual Template with CCA. In *Proceedings of the 2019 3rd International Conference on Innovation in Artificial Intelligence* (Suzhou, China) *(ICIAI 2019)*. Association for Computing Machinery, New York, NY, USA, 6–10. https://doi.org/10.1145/3319921.3319925

[41] R Martinez Orellana, Burak Erem, and Dana H Brooks. 2013. Time invariant multi electrode averaging for biomedical signals. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 1242–1246.

[42] B. Rebsamen, E. Burdet, Q. Zeng, H. Zhang, M. Ang, C. L. Teo, C. Guan, and C. Laugier. 2008. Hybrid P300 and Mu-Beta Brain Computer Interface to Operate a Brain Controlled Wheelchair. In *Proceedings of the 2nd International Convention on Rehabilitation Engineering & Assistive Technology* (Bangkok, Thailand) *(iCREATe '08)*. Singapore Therapeutic, Assistive & Rehabilitative Technologies (START) Centre, Midview City, SGP, 51–55.

[43] Maurice Rekrut, Tobias Jungbluth, Jan Alexandersson, and Antonio Krüger. 2021. Spinning Icons: Introducing a Novel SSVEP-BCI Paradigm Based on Rotation. In *26th International Conference on Intelligent User Interfaces* (College Station, TX, USA) *(IUI '21)*. Association for Computing Machinery, New York, NY, USA, 234–243. https://doi.org/10.1145/3397481.3450646

[44] Zahra Rezazadeh and Ali Sheikhani. 2017. An SSVEP-Based Brain-Computer Interface to Navigate in a Virtual Home. In *Proceedings of the 7th International Conference on Bioscience, Biochemistry and Bioinformatics* (Bangkok, Thailand) *(ICBBB '17)*. Association for Computing Machinery, New York, NY, USA, 22–27. https://doi.org/10.1145/3051166.3051171

[45] Dieter Schmalstieg, Alexander Bornik, Gernot Müller-Putz, and Gert Pfurtscheller. 2010. Gaze-Directed Ubiquitous Interaction Using a Brain-Computer Interface. In *Proceedings of the 1st Augmented Human International Conference* (Megève, France) *(AH '10)*. Association for Computing Machinery, New York, NY, USA, Article 4, 5 pages. https://doi.org/10.1145/1785455.1785459

[46] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3549–3557.

[47] Deep Soni, Nitesh Singh Malan, and Shiru Sharma. 2019. CCA Model with Training Approach to Improve Recognition Rate of SSVEP in Real Time. In *Proceedings of the 2019 3rd International Conference on Artificial Intelligence and Virtual Reality* (Singapore, Singapore) *(AIVR 2019)*. Association for Computing Machinery, New York, NY, USA, 56–59. https://doi.org/10.1145/3348488.3348498

[48] Ramesh Srinivasan, F Alouani Bibi, and Paul L Nunez. 2006. Steady-state visual evoked potentials: distributed local sources and wave-like dynamics are sensitive to flicker frequency. *Brain topography* 18, 3 (2006), 167–187. https://doi.org/10.1007/s10548-006-0267-4

[49] Piotr Stawicki, Felix Gembler, Cheuk Yin Chan, Mihaly Benda, Aya Rezeika, Abdul Saboor, Roland Grichnik, and Ivan Volosyak. 2018. Investigating Spatial Awareness within an SSVEP-based BCI in Virtual Reality. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 615–618. https://doi.org/10.1109/SMC.2018.00113

[50] Piotr Stawicki, Felix Gembler, Cheuk Yin Chan, Mihaly Benda, Aya Rezeika, Abdul Saboor, Roland Grichnik, and Ivan Volosyak. 2018. SSVEP-Based BCI in Virtual Reality - Control of a Vacuum Cleaner Robot. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 534–537. https://doi.org/10.1109/SMC.2018.00749

[51] Piotr Stawicki, Felix Gembler, Aya Rezeika, and Ivan Volosyak. 2017. A Novel Hybrid Mental Spelling Application Based on Eye Tracking and SSVEP-Based BCI. *Brain Sciences* 7, 4 (2017). https://doi.org/10.3390/brainsci7040035

[52] Piotr Stawicki, Felix Gembler, and Ivan Volosyak. 2015. Evaluation of Suitable Frequency Differences in SSVEP-Based BCIs. In *Symbiotic Interaction*, Benjamin Blankertz, Giulio Jacucci, Luciano Gamberini, Anna Spagnolli, and Jonathan Freeman (Eds.). Springer International Publishing, Cham, 159–165.

[53] Fei Teng, Aik Min Choong, Scott Gustafson, Dwight Waddell, Pamela Lawhead, and Yixin Chen. 2010. Steady State Visual Evoked Potentials by Dual Sine Waves. In *Proceedings of the 48th Annual Southeast Regional Conference* (Oxford, Mississippi) *(ACM SE '10)*. Association for Computing Machinery, New York, NY, USA, Article 50, 4 pages. https://doi.org/10.1145/1900008.1900077

[54] D. A. Todd, P. J. McCullagh, M. D. Mulvenna, and G. Lightbody. 2012. Investigating the Use of Brain-Computer Interaction to Facilitate Creativity. In *Proceedings of the 3rd Augmented Human International Conference* (Megève, France) *(AH '12)*. Association for Computing Machinery, New York, NY, USA, Article 19, 8 pages. https://doi.org/10.1145/2160125.2160144

[55] Pablo Martinez Vasquez, Hovagim Bakardjian, Montserrat Vallverdu, and Andrezj Cichocki. 2008. Fast multi-command SSVEP brain machine interface without training. In *International Conference on Artificial Neural Networks*. Springer, 300–307.

[56] François-Benoît Vialatte, Monique Maurice, Justin Dauwels, and Andrzej Cichocki. 2010. Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives. *Progress in Neurobiology* 90, 4 (2010), 418–438. https://doi.org/10.1016/j.pneurobio.2009.11.005

[57] Ivan Volosyak, Hubert Cecotti, and Axel Gräser. 2009. Impact of frequency selection on LCD screens for SSVEP based brain-computer interfaces. In *International Work-Conference on Artificial Neural Networks*. Springer, 706–713.

[58] Lisa-Marie Vortmann and Felix Putze. 2020. Attention-Aware Brain Computer Interface to Avoid Distractions in Augmented Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382889

[59] Yao Wang, Kun Li, Xiang Zhang, Jinhai Wang, and Ran Wei. 2020. Research on the Application of Augmented Reality in SSVEP-BCI. In *Proceedings of the 2020 6th International Conference on Computing and Artificial Intelligence* (Tianjin, China) *(ICCAI '20)*. Association for Computing Machinery, New York, NY, USA, 505–509. https://doi.org/10.1145/3404555.3404587

[60] Yijun Wang, Ruiping Wang, Xiaorong Gao, Bo Hong, and Shangkai Gao. 2006. A practical VEP-based brain-computer interface. *IEEE Transactions on neural systems and rehabilitation engineering* 14, 2 (2006), 234–240.

[61] Yijun Wang, Zhiguang Zhang, Xiaorong Gao, and Shangkai Gao. 2004. Lead selection for SSVEP-based brain-computer interface. In *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 2. 4507–4510. https://doi.org/10.1109/IEMBS.2004.1404252

[62] Dong-Ok Won, Hai Hong Zhang, Cuntai Guan, and Seong-Whan Lee. 2014. A BCI speller based on SSVEP using high frequency stimuli design. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 1068–1071. https://doi.org/10.1109/SMC.2014.6974055

[63] Yongcheng Wu, Guanghua Xu, Yifan Wu, Bo Wang, Nan Duan, Liang Zeng, Zezhen Han, and Sicong Zhang. 2020. An Asynchronous Detection Algorithm for SSVEP-Based BCI Using Gradient Boosting Decision Tree. In *Proceedings of the 2020 9th International Conference on Computing and Pattern Recognition* (Xiamen, China) *(ICCPR 2020)*. Association for Computing Machinery, New York, NY, USA, 101–105. https://doi.org/10.1145/3436369.3437405

[64] Jun Xie, Guanghua Xu, Jing Wang, Min Li, Chengcheng Han, and Yaguang Jia. 2016. Effects of Mental Load and Fatigue on Steady-State Evoked Potential Based Brain Computer Interface Tasks: A Comparison of Periodic Flickering and Motion-Reversal Based Visual Attention. *PLOS ONE* 11, 9 (09 2016), 1–15. https://doi.org/10.1371/journal.pone.0163426

[65] Jun Xie, Guanghua Xu, Jing Wang, Feng Zhang, and Yizhuo Zhang. 2012. Steady-State Motion Visual Evoked Potentials Produced by Oscillating Newton's Rings: Implications for Brain-Computer Interfaces. *PLOS ONE* 7, 6 (06 2012), 1–8. https://doi.org/10.1371/journal.pone.0039707

[66] Danhua Zhu, Jordi Bieger, Gary Garcia Molina, and Ronald M Aarts. 2010. A survey of stimulation methods used in SSVEP-based BCIs. *Computational intelligence and neuroscience* 2010 (2010).