# Parameter Optimization for Brain-Computer Interfaces based on Visual Evoked Potentials

---

Felix Gembler

# Parameter Optimization for Brain-Computer Interfaces based on Visual Evoked Potentials

zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEUR (Dr.-Ing.)

der Technischen Fakultät
der Universität Bielefeld

Dissertation

von

Felix Gembler

| | |
|---|---|
| Referent: | Prof. Dr.-Ing. Ulrich Rückert |
| Korreferent: | Prof. Dr.-Ing. Ivan Volosyak |
| Korreferentin: | Jun.-Prof. Dr. Karolin Schäfer |

Tag der mündlichen Prüfung: 29.10.2020

Gedruckt auf alterungsbeständigem Papier (ISO 9706)

# Acknowledgment

First of all, I would like to thank my research supervisors Prof. Dr.-Ing Ivan Volosyak and Prof. Dr.-Ing Ulrich Rückert, for their continuous support of my work, their motivation and their guidance throughout this project.

I am grateful to my Hochschule Rhein-Waal colleagues, most of all, the wonderful team members of the BCI-Lab: For helpful ideas and contributions to the initial software, Piotr Stawicki. For proofreading and valuable comma placement, Aya Rezeika. For the contributions to the dictionary implementation, Abdul Saboor. For providing helpful feedback and aiding in rubber duck debugging, Mihaly Benda.

I thank all student assistants of the BCI-Lab for their assistance in conducting EEG experiments. I am further grateful for the staff of the Franziskushaus and Bruderschaft zu unserer lieben Frau for assistance in the recruitment of participants.

Many thanks go to all the participants who tested our BCI systems. I want to especially thank Kathrin Lemler from the University of Cologne, who volunteered to test one of our prototypes.

Besides, I thank Jun.-Prof. Dr. Karolin Schäfer and Prof. Dr. Jens Boenisch from the University of Cologne for continuing our experiments with patients.

Last but not least, I am forever grateful to my beloved family, my sister and my parents for their continuous moral support, encouragement, and faith in me.

*Felix Gembler*

# Contents

# List of Figures

# List of Tables

# Abstract

For people who are physically unable to communicate with their fellow human beings due to severe disabilities, technical communication aids can be a life enrichment. Communication tools can be realized with so-called brain-computer interfaces (BCIs), which provide a connection between the brain and the computer and can be controlled without the activation of the peripheral nervous system. The electrical brain activity is recorded, usually non-invasively, by means of an electroencephalogram (EEG). BCIs analyze the collected EEG data in real-time and convert them into output signals allowing hands-free control of various kinds of applications such as mental typewriters.

One of the control paradigms used to realize BCIs is based on visual evoked potentials (VEPs), which appear in the visual cortex of the brain when visual stimuli are perceived. An example of such stimuli is flickering target objects on a computer screen, each flashing with a specific frequency. By detecting the VEPs, the BCI can determine the target on which the user is focusing. In spelling applications, these targets represent letters; the user can spell a word or sentence just by looking at the corresponding stimulus.

In several studies, a high variation in BCI accuracy across users has been observed; not all users did achieve reliable control over the system. A significant problem in BCI research is that EEG data cannot be interpreted reliably for all users. In spelling applications, the system might output wrong letters too frequently, which makes effective communication difficult. For other applications, such as wheelchair control, faulty classifications should be avoided entirely.

An essential goal in the field of research is, therefore, to improve the accuracy of the classification. One way to achieve this goal is to customize critical parameters to the user.

In this work, factors that impact the performance of VEP-based BCIs were investigated. These factors include parameters and settings of the user interface and the classification, such as the number of targets and the duration that a stimulus needs to be fixated until the corresponding command is executed. Furthermore, demographic differences such as age and gender and their relation to BCI performance were analyzed. To this end, several studies - each dedicated to one or several of these factors - were conducted. The results of these studies indicated that user age and the number of targets of the graphical user interface have a high impact on classification accuracy.

Based on these findings, a robust BCI application was developed, a spelling application that determines personalized key parameters. This application enables a more accurate BCI control, as the BCI is tailored to the respective user. Moreover, the software allows non-specialists to set up the system. The latter is an essential point in terms of usability in daily life; it enables nursing staff or family members to adjust the necessary system settings with little effort.

# Zusammenfassung

Für Menschen, die wegen schweren Erkrankungen körperlich nicht mehr in der Lage sind mit ihren Mitmenschen zu kommunizieren, können technische Kommunikationshilfen eine Lebensbereicherung sein. Eine solche Kommunikationshilfe kann mit sogenannten Brain-Computer-Interfaces (BCIs, deutsch Gehirn-Computer-Schnittstellen) realisiert werden. BCIs ermöglichen eine Verbindung zwischen dem Gehirn und dem Computer und können ohne eine Aktivierung des peripheren Nervensystems bedient werden. Hierzu wird die elektrische Gehirnaktivität aufgezeichnet. Dies kann zum Beispiel nicht-invasiv mittels eines Elektroenzephalogramms (EEG) erfolgen. BCIs analysieren die aufgenommenen EEG-Daten in Echtzeit und wandeln sie in Ausgangssignale um. Auf diese Weise können diverse Applikationen, wie zum Beispiel Schreibprogramme, ohne die Nutzung der Hände angesteuert werden.

Zu den Paradigmen, mit denen BCIs realisiert werden können, zählen visuell evozierte Potentiale (VEP). Diese elektrischen Potentiale treten im visuellen Cortex (auch Sehrinde) des Gehirns auf, wenn visuelle Stimuli wahrgenommen werden. Ein Beispiel solcher Stimuli sind mit verschiedenen konstanten Frequenzen flackernde Boxen auf einem Computermonitor. Durch die Detektion der VEPs kann das BCI feststellen, welche Box angesehen wird. Bei Schreibprogrammen repräsentieren diese Boxen Buchstaben, so dass Personen, die das System nutzen, nur durch das Ansehen der entsprechenden Box ein Wort oder einen Satz buchstabieren können.

In einigen Studien wurde eine hohe Variabilität der Klassifikationsgenauigkeit zwischen den Nutzenden beobachtet; dabei konnte in einigen Fällen keine zuverlässige Kontrolle über das System erreicht werden. Ein wesentliches Problem der BCI-Technologie liegt also darin, dass die EEG-Daten nicht immer zuverlässig interpretiert werden können. Das führt dazu, dass etwa bei Schreibprogrammen häufig ungewünschte Buchstaben ausgewählt werden, was die Kommunikation erschwert. Für andere Applikationen, beispielsweise eine Rollstuhlansteuerung, sollten fehlerhafte Klassifikationen komplett vermieden werden.

Ein wesentliches Ziel ist es daher die Genauigkeit der BCI-Klassifikation zu verbessern. Ein Ansatz dieses Ziel zu erreichen ist die individuelle Anpassung wichtiger BCI-Parameter an den Nutzer oder die Nutzerin.

In dieser Arbeit wurden unterschiedliche Faktoren, die die Performanz VEP-basierter BCIs beeinflussen können, untersucht. Hierzu zählen Parameter und Einstellungen der Nutzerschnittstelle und der Verarbeitungsalgorithmen, etwa die Anzahl der zu unterscheidenden Ziele und die Dauer, die ein Stimulus fixiert werden muss, damit das entsprechende Kommando ausgeführt wird. Darüber hinaus wurde die BCI Performanz in Bezug auf demographische Unterschiede, wie das Alter und das Geschlecht, analysiert. Mehrere Studien wurden durchgeführt, um diese Faktoren zu untersuchen. Die Ergebnisse dieser Studien deuten darauf hin, dass das Alter der Nutzerinnen und Nutzer und die Anzahl der BCI-Ziele

der graphischen Benutzeroberfläche einen hohen Einfluss auf die Klassifikationsgenauigkeit haben.

Auf Basis dieser Erkenntnisse wurde eine robuste BCI-Anwendung entwickelt; ein Schreibprogramm bei dem wichtige Parameter auf die Nutzerin oder den Nutzer angepasst werden. Zum einen ermöglicht diese Anwendung eine exaktere BCI Ansteuerung. Ferner ermöglicht sie auch Laien die Einstellung des Systems. Letzteres ist ein wichtiger Punkt, denn so können die notwendigen Einstellungen beim alltäglichen Gebrauch des Systems mit nur wenig Aufwand vom Pflegepersonal oder von Familienmitgliedern vorgenommen werden.

# 1 Introduction

*"Tech changed my life, without it I would be mute - I would still be the same chatterbox, but nobody would notice"*
— Kathrin Lemler, *Board member of the Gesellschaft für Unterstützte Kommunikation e.V.*

Several assistive technologies are being developed to enhance the life quality of people with severe disabilities. Kathrin Lemler from the University of Cologne, who suffers from infantile cerebral palsy due to a lesion in the central nervous system in early childhood, used to communicate with a simple letter board. Nowadays, she uses an eye-tracking device in her daily life; the input modality of this device is based on eye movements.

Another way to realize assistive technologies are brain-computer interfaces (BCIs), which use recorded brain signals as an input modality. A variation of these systems can be realized by employing certain visual stimuli, specific flickering patterns that elicit visual evoked potentials (VEPs) in the brain. Using constant frequencies, which evoke steady-state VEPs (SSVEPs), or pseudorandom code patterns, which evoke code-modulated VEPs (c-VEPs), various types of applications, such as spelling interfaces [1] and control applications for a prosthesis [2] or navigation [3] can be operated.

There is a growing research interest in BCI technologies [4]; the focus of this thesis lies exclusively on VEP-based BCIs, which represent a standard BCI paradigm and have become quite common over time (see Figure 1.1).

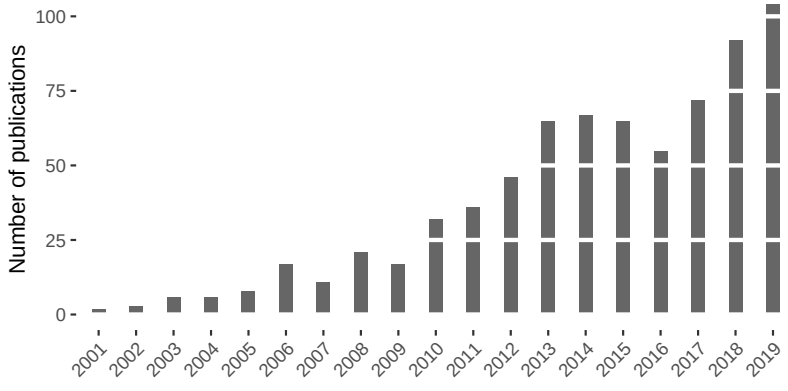



**Figure 1.1:** Publications from 2001 - 2019 related to BCIs based on steady-state visual evoked potentials (SSVEPs) or code-modulated visual evoked potentials (c-VEPs). Displayed are the number of articles identified from PubMed search results for the search term "ssvep or c-vep or cvep".

## 1.1 Motivation and Problem Statement

Over the last 20 years, many studies related to VEP-based BCIs have been conducted. Generally, promising results have been achieved. VEP-based BCIs yield high spelling speeds and have been successfully tested with disabled users [5–7].

On the other hand, cases where the BCI system was not able to interpret the user's intent with sufficient accuracy have been reported repeatedly [1, 8, 9].

This issue - commonly referred to as BCI illiteracy (or BCI deficiency) - is a significant problem in BCI research. It may even reduce the public interest in the technology as it hinders its use in practical scenarios and complicates commercial exploitation.

In many BCI studies relatively small sample groups with a low mean age are recruited. Only a few BCI field studies were conducted in the last two decades, which repeatedly reported cases of BCI illiteracy. Especially elderly users were affected. The relationship between age and BCI performance has yet to be carefully studied.

The occurrence of BCI illiteracy could be reduced if system parameters are identified and adjusted carefully for each user. To this end, the following questions need to be addressed: What are the main factors contributing to sufficient BCI control? To what extent does the number of targets, the frequency choice, and the length of the classification time window affect BCI performance? Do age and gender impact the BCI performance?

In real-life scenarios, the setup of BCI parameters needs to be applied by non-experts (e.g., family members or caregivers). A primary goal of this thesis was, therefore, the development of a robust BCI application that can be customized by non-experts while simultaneously yielding a high BCI literacy rate (i.e., a high percentage of users that achieve reliable control).

To achieve this goal, this thesis addresses the following principal aims:

- to investigate age-related differences in VEP BCI performance;
- to explore key system parameters and their impact on BCI performance and BCI illiteracy;
- to implement auto-calibration methods;
- to improve signal classification methods; and
- to synthesize the above finding to propose a robust BCI for all that can be set up by non-experts.

## 1.2 Publications

This thesis presents and summarizes research results and findings achieved during my work as a research assistant at the BCI-Lab Kleve, Rhine-Waal University of Applied Sciences. Some results, ideas, and figures of the following published scientific papers were incorporated:

- I. Volosyak, F. Gembler, and P. Stawicki. 'Age-Related Differences in SSVEP-Based BCI Performance'. In: *Neurocomputing* 250 (2017), pp. 57–64. DOI: 10.1016/j.neucom.2016.08.121.
- F. Gembler, P. Stawicki, and I. Volosyak. 'Autonomous Parameter Adjustment for SSVEP-Based BCIs with a Novel BCI Wizard'. In: *Frontiers in Neuroscience* 9 (Dec. 2015). DOI: 10.3389/fnins.2015.00474.
- F. Gembler, P. Stawicki, and I. Volosyak. 'Suitable Number of Visual Stimuli for SSVEP-Based BCI Spelling Applications'. In: *Advances in Computational Intelligence: 14th International Work-Conference on Artificial Neural Networks, IWANN 2017, Cadiz, Spain, June 14-16, 2017, Proceedings, Part II*. Ed. by I. Rojas, G. Joya, and A. Catala. Cham: Springer International Publishing, 2017, pp. 441–452. DOI: 10.1007/978-3-319-59147-6_38.
- F. Gembler, P. Stawicki, and I. Volosyak. 'Exploring the Possibilities and Limitations of Multitarget SSVEP-Based BCI Applications'. In: *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the the IEEE Engineering in Medicine and Biology Society (EMBC)*. Orlando, FL, USA, 2016, pp. 1488–1491. DOI: 10.1109/EMBC.2016.7590991.
- F. Gembler and I. Volosyak. 'A Novel Dictionary-Driven Mental Spelling Application Based on Code-Modulated Visual Evoked Potentials'. In: *Computers* 8.2 (2019). DOI: 10.3390/computers8020033.
- F. Gembler, P. Stawicki, A. Rezeika, and I. Volosyak. 'A Comparison of cVEP-Based BCI-Performance Between Different Age Groups'. en. In: *Advances in Computational Intelligence*. Ed. by I. Rojas, G. Joya, and A. Catala. Vol. 11506. Cham: Springer International Publishing, 2019, pp. 394–405. DOI: 10.1007/978-3-030-20521-8_33.
- F. Gembler, P. Stawicki, A. Saboor, and I. Volosyak. 'Dynamic Time Window Mechanism for Time Synchronous VEP-Based BCIs—Performance Evaluation with a Dictionary-Supported BCI Speller Employing SSVEP and c-VEP'. en. In: *PLOS ONE* 14.6 (June 2019). Ed. by Z. Wang, e0218177. DOI: 10.1371/journal.pone.0218177.

- F. Gembler, M. Benda, A. Saboor, and I. Volosyak. 'A Multi-Target c-VEP-Based BCI Speller Utilizing n-Gram Word Prediction and Filter Bank Classification'. In: *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). Oct. 2019, pp. 2719–2724. DOI: 10.1109/SMC.2019.8914235.

In case one of the publications listed above was fundamental to a section, it is stated at the beginning of that section.

## 1.3 Thesis Structure

This thesis is structured as follows. Chapter 2 provides definitions, nomenclature, and historical backgrounds related to the field of BCIs. Different types of spelling applications based on the VEP-paradigm are reviewed in section 2.6. The evaluation metrics used to assess BCI performance are defined in section 2.7.

Chapter 3 describes the components required for the implementation of VEP-based BCIs: Methods of stimulus presentation are described in sections 3.1 and 3.2, the implementation of stimuli and classification methods are targeted in sections 3.3, 3.4, and 3.5.

Chapters 4 and 5 summarize the studies listed in section 1.2, which were conducted to investigate the primary research questions of this thesis. While chapter 4 focuses on parameters for the SSVEP paradigm, Chapter 5 addresses the c-VEP paradigm. In both chapters, studies conducted to investigate key system parameters and automated calibration are presented. As the methods and materials of these studies are similar, they are provided in detail for the first study summarized in the respective chapter. Impacts of user age on performance are investigated in sections 4.3 and 5.2, methods of auto-calibration are demonstrated in sections 4.4 and 5.3, and the optimal number of BCI targets is discussed in sections 4.5 and 5.4. Additionally, the SSVEP and c-VEP paradigms are compared in section 5.3, addressing user-friendliness and overall system speed.

Finally, chapter 6 summarizes the thesis, outlines the conclusions, and provides recommendations for further research.

# Brain-Computer Interfaces 2

The invention of the electroencephalography (EEG) in the early twentieth century led to the realization of communication via brain activity. The expression 'brain-computer interface' for a system that links the human brain to a computer was initially introduced by Vidal in 1973 [10]. A few years later, in 1977, the first BCI application, a cursor-object on a computer screen that could be controlled by means of EEG analysis, was presented [11]. Since then, great advances have been made; many laboratories began researching BCIs, exploring various control paradigms and applications while forming a standard knowledge structure (taxonomy).

This chapter provides a brief introduction to BCIs (section 2.1) and the EEG recording technique (section 2.2). Following that, the major BCI control paradigms, the general BCI framework, and typical applications are presented (sections 2.3, 2.4 and 2.5). Then, the focus is put on VEP-based spelling applications (section 2.6). After that, the evaluation metrics used in this thesis are described. In particular, the classification accuracy, the information transfer rate (ITR) in bit per minute (bpm), and the output characters per minute (OCM), are introduced (section 2.7). The chapter closes with a discussion of BCI illiteracy (section 2.8).

## 2.1 Definition

In 2002, Wolpaw et al. [12] shaped the definition that is now recognized in the research field by describing the BCI as a system that enables its user to interact with his or her environment, without the use of peripheral nerves and muscles. Figure 2.1 illustrates this concept: The BCI interprets recorded patterns of the user's brain activity employing physiological principals related to specific cognitive tasks; the generated commands can be used to control various kinds of applications.

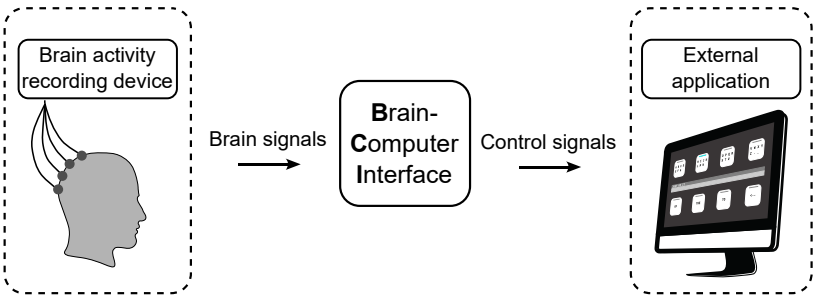



**Figure 2.1:** Brain-computer interface. The system translates recorded brain signals into control signals for an external application.

Among the most popular fields of applications for BCIs are communication tools for patients suffering from motor neuron disease (MND) [13]. MNDs such as amyotrophic lateral sclerosis (ALS) disrupt neurological networks affecting the brain's communication with other organs and the motor control of the muscles. Patients suffering from locked-in syndrome (LIS), brainstem stroke, brain or spinal cord injury, cerebral palsy, muscular dystrophies, or multiple sclerosis eventually lose voluntary control over skeletal muscles and the tongue, which leads to difficulties in exchanging of information. These patients can benefit from alternative communication systems. For example, communication tools based on eye-tracking allow the control of a virtual keyboard by tracking the movement of the pupils; if the gaze lingers on the desired letter, the system produces the associated output command [14]. Unfortunately, these kinds of applications are not suitable for persons who cannot control ocular movements or who experience uncontrollable head movements [15]. However, these patients might still be able to communicate via brain signals using a BCI.

Researchers have categorized BCIs according to several characteristics; the following classes are typically used to describe certain aspects of the system (e.g., [13]):

**Dependent/ independent:**
Dependent BCIs require control over peripheral nerves and muscles; for example, gaze dependent systems require control over extraocular muscles. In contrast, independent BCIs only rely on brain activity; they do not depend on any muscle activity. Therefore, independent systems are suitable for patients who lost control over their eye muscles, such as late-stage ALS patients.

**Exogenous/endogenous:**
Exogenous BCIs require an external stimulus (for example, a visual flickering pattern). Endogenous BCIs, on the other hand, are based on brain patterns that are voluntarily modulated by the user's imagination; they do not require any external stimuli. For these kinds of systems, the user needs to learn how to modulate his or her brain signals in such a way that the BCI interprets them as desired.

**Synchronous/asynchronous:**
In synchronous BCIs, the system determines the timings of the control, i.e., the user does not influence the time a system output is generated. Conversely, in asynchronous BCIs, users control the timing of produced commands, which results in a more natural interaction between user and system.

**Invasive/noninvasive:**
Systems that do not require surgery are referred to as non-

invasive BCIs. An example of noninvasive BCIs are systems based on EEG, which record brain activity via electrodes placed on the scalp. On the other hand, systems that do require surgery are referred to as invasive BCIs. For example, systems based on electrocorticography (ECoG) require direct access to the brain tissue and, thus, an invasive opening through the skull. The brain signals are recorded via electrodes implanted under the scalp. Because of the involved costs and risks, invasive BCIs are rarely used in research with humans.

The various methods to monitor brain activity include magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), near-infrared spectroscopy (NIRS), ECoG, and EEG; the latter approach, EEG, is the most practical method, as it is non-invasive, portable, comparably inexpensive, and easy to apply [16].

## 2.2 Electroencephalography

Since 1924, when physiologist Hans Berger [17] recorded the first human EEG, the procedure has become a widely adopted method in research and medicine, which significantly advanced the understanding of the brain's functional architecture. Because of its practicability, widespread availability, and high temporal resolution, EEG is by far the most applied monitoring method in BCI research [16].

EEG measures the electric potential generated by the summed synchronous activity from millions of neurons over time. The voltage fluctuations of these summed patterns are still small, ranging from -100 to 100 $\mu$V [18]. For this reason, EEG measurements are sensitive to external electrical noise; the recorded signals need to be amplified.

The basic setup of the recording system consists of electrodes, amplifiers, analog-to-digital converter (ADC), and a recording device. The electrodes acquire the signal from the surface of the head; the amplifiers enlarge the signal amplitudes for accurate digitization by the ADC; the recording device (typically a personal computer) stores and displays the collected data. EEG measures the potential difference between a signal and a reference electrode over time. The voltages are measured against an additional electrode, the common ground, so that artifacts from power sources, common to all signal electrodes and the reference electrode are reduced. While three electrodes (ground, reference, and signal electrode) are the minimal configuration for EEG, multi-channel configurations

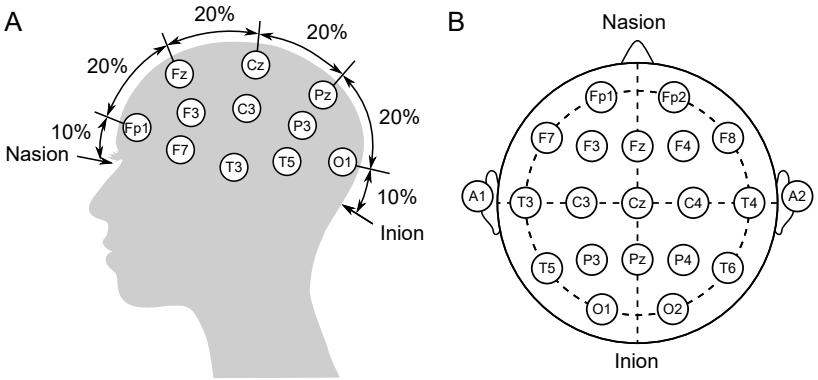



**Figure 2.2:** Standard 10-20 electrode placement [22]. (A) Left side of head. (B) Top of head scheme. The position labels are composed of letters specifying the lobe and numbers determining the hemisphere location. The letters A, F, Fp, T, C, P, and O denote ear, frontal, frontal polar, temporal, central, parietal, and occipital lobes, respectively. Even numbers correspond to electrodes placed on the right hemisphere; odd numbers correspond to electrodes placed on the left hemisphere; A lower case 'z' (i.e., zero) refers to electrodes placed on the middle line.

can comprise of many more additional signal electrodes. Some applications employ more than 100 signal electrodes [19].

An advantage of EEG is the high temporal resolution; the EEG activity is scanned on the level of milliseconds, depending on the sampling frequency (sampling frequencies between 100 and 1000 Hz are common in BCI research). A disadvantage of EEG is the low spatial resolution on the scalp in comparison to other methods, which is caused by various resistive layers (especially the skull) between the electrical sources within the brain and the recording electrodes. These layers induce a smearing effect (also called volume-conduction effect) [20].

For suitable signal quality impedances between sensors and scalp need to be lowered (typically below 5 kΩ is recommended, depending on the type of electrode [21]). In this regard, a non-abrasive electrode gel needs to be applied to the scalp, which serves as a conductive pass between skin and electrodes. Electrodes can be classified as active or passive electrodes. Active electrodes are less affected by external noise, such as cable movements, as they use pre-amplification very close to the skin at the recording sites.

As the preparation and cleanup procedure when using gel-based electrodes is quite cumbersome, several alternatives have been tested. For example, water-based sensors (small pellet electrodes rolled in cotton soaked with water) may simplify the setup [23]. Moreover, dry electrodes which are integrated into the cap or affixed on top of the scalp, do not require any gel at all [24, 25]. Unfortunately, the signal quality might be considerably lower with these types of electrodes [24].

Figure 2.2 shows the international standard 10-20 electrode system montage, which was developed with the goal of reproducible measurements [22]. The 10-20 system standardizes 21 positions for electrode placement using the nasion (at the top of the nose) and the inion (at the bony lump at the base of the skull) as reference

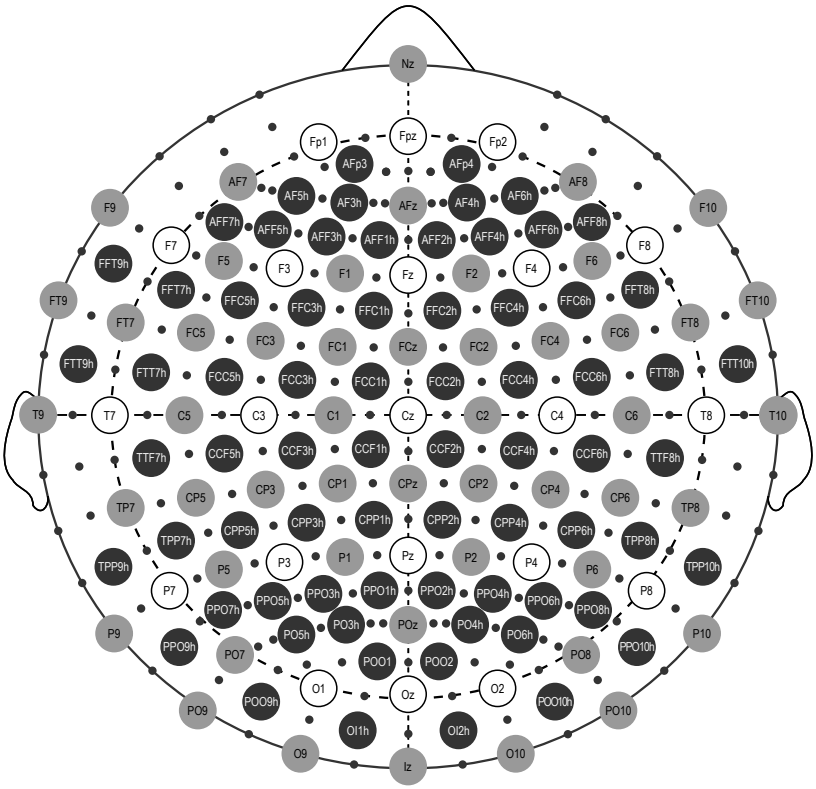



**Figure 2.3:** The 10-5 electrode system [19]. Electrode locations of the standard 10-20 system are indicated by white circles, additional electrode locations of the 10-10 system are indicated by gray circles, and additional electrode locations of the 10-5 system are indicated by black dots and circles; the black circles indicate a selection of locations suitable for a 128 channel EEG system.

points. The electrodes are labeled according to the distance between these references, which is divided into intervals of 10% and 20%. Each location corresponds to a position label composed of a letter and a number. The letter refers to the brain region: The letter A refers to the electrode placement on the mastoid process (the bony prominence behind the ear), C to the central, P to the parietal, F to the frontal, Fp to the frontal polar, and O to the occipital area. The number refers to the lateral distance to the mid-line. The positions on the mid-line have no number; instead, the letter z (for zero) is appended. Odd numbers refer to the left and even numbers to the right side; a higher number indicates a greater distance to the mid-line.

Over time, the 10-20 system has been extended to the 10-10 and 10-5 systems, allowing up to 74 and 345 electrode positions, respectively [19]. Figure 2.3 shows the electrode locations of the 10-10 system (marked as gray circles) and of the 10-5 systems (marked as black circles and dots). The 10-10 system uses letter combinations to label the additional positions, e.g., the electrodes between P and O are labeled PO. The 10-5 system appends the letter 'h' (half) for electrode positions halfway between two existing locations.

Depending on the type of BCI, the amplitude, latency, or frequency spectrum of the sampled raw EEG data can be investigated. The data can be decomposed into its constituent frequencies via Fourier transformation. Certain frequencies may be more dominant in the EEG; several frequency ranges (frequency bands) have been defined

to categorize the brain waves and associated characteristics [18, 26]: the delta band (< 4 Hz, associated with deep sleep), the theta band (4-7 Hz, associated with drowsiness), the alpha band (8-12 Hz, associated with relaxed states and closing the eyes, recorded over the occipital region), the mu rhythms (8-12 Hz, associated with motor activities, recorded over the sensorimotor region), the beta rhythms (12-30 Hz, associated with motor activities, recorded over frontal and central regions), and the gamma rhythms (30-100 Hz, associated with motor activities and meditative states). The exact boundaries of the bands are not consistent in literature; for example, the alpha band is sometimes reported from 8 to 15 Hz.

## 2.3 The Three Major BCI Paradigms

The BCI research field can be split into several sub-categories on the basis of the used control paradigm. Each of these paradigms is based on a specific brain potential occurring in the EEG. In the following, the three most common BCI control paradigms are described: The MI-based paradigm, which is based on sensorimotor rhythms (SMRs), the P300 paradigm which is based on event-related potentials (ERPs) and the VEP paradigm, which includes the SSVEP and the c-VEP control signals. Figure 2.4 provides example applications for each of these paradigms.

### The Motor Imagery Paradigm

SMRs are brain waves recorded over the sensorimotor cortex [29]. The amplitudes of these rhythms change when a person is moving, but also when preparing or imagining the movement [30]. SMR rhythms can be characterized by frequency bands; most commonly used are the $\mu$ rhythm and the $\beta$ rhythm [31].

Each body part corresponds to a specific section of the brain that controls the movement of that part. For example, the right hand movement corresponds to motor cortex activity in the contralateral hemisphere; it is, therefore, possible to distinguish the brain activity corresponding to the movement of the left and the right hand [30]. When not involved in a motor task, the measured signal at the corresponding motor area in the brain shows synchronized activity in the $\mu$ band; this activity is referred to as event-related synchronization (ERS). On the other hand, during a motor task, for example, a hand movement, a desynchronization activity can be observed, where the amplitude of the $\mu$ rhythm decreases; this activity is referred to as event-related desynchronization (ERD). ERS and ERD are also observed when the movement is only imagined rather than executed [32]. For that reason, SMR activity

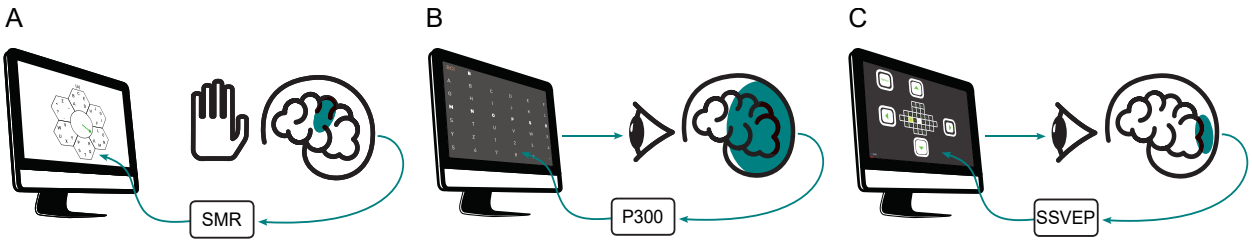



**Figure 2.4:** The three major BCI paradigms. (A) The motor imagery (MI) paradigm. The imagination of the movement of a limb, e.g., the imaginary movement of the hand, as in the picture, generates a sensorimotor rhythm (SMR) signal which can be recorded over the motor cortex. The figure shows the graphical user interface (GUI) of the Hex-o-Spell system [27], where the centrally positioned arrow can be rotated clockwise by imagining right hand movement. This control mechanism allows the selection of hexagons representing groups of letters or, in a second step, individual letters. (B) The P300 paradigm. The BCI presents a character grid to the user with various rows and columns highlighted. The infrequent event, where the highlighted column or row contains the desired target character, elicits the P300 peak roughly 300 ms after stimulus onset. The system then matches the detected P300 to the target character. The figure shows the Farwell and Donchin speller [28], which presents a $6 \times 6$ matrix of symbols, including all 26 letters of the alphabet as well as 10 additional symbols. (C) The steady-state visual evoked potential (SSVEP) paradigm. The GUI presents different stimulation frequencies to the user. By gazing at one of them, SSVEPs are elicited in the visual cortex of the brain. The SSVEP-BCI analyzes the data measured via electroencephalography (EEG) and produces an output command. The figure shows the Bremen BCI [1], which employs five stimulation frequencies. Figure adapted from Rezeika et al. [13].

is used as input for MI-based BCIs. These systems interpret brain activity, which the user modulates by imagining limb movements or by performing similar cognitive tasks. Figure 2.4 A shows an MI-based spelling application, the Hex-o-Spell system [27], which is controlled by imagining right hand movements.

MI-based BCIs are endogenous and independent, as they can be operated with brain activity only and do not require external stimuli. These systems are, therefore, suitable for patients with affected sensory organs. Successful tests with ALS patients have been conducted with MI-based BCIs [33].

MI applications are typically synchronous as they specify fixed time windows where the user needs to imagine the movement after a cue signal [34]. It should be noted that MI-based BCIs typically require more training time than other paradigms, as the user needs to learn how to modulate the SMR rhythms in several sessions [29, 33]. However, the use of modern machine learning methods has lead to a strong decrease in calibration time from more than 50 hours to less than half an hour [35].

## The P300 Paradigm

ERPs are electrocortical signals, which are detectable via EEG and occur as a response to a sensory (visual, auditory, or tactile) or psychological event. The type of the stimulus event determines the characteristics (i.e., the specific fixed time delay, location, and amplitude) of the evoked ERPs. Among the most researched ERPs in BCI literature is the P300.

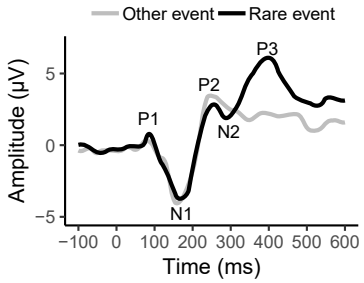



**Figure 2.5:** P300 component of an event-related potential. The trial with the P300 component (elicited by a rare event) shows a large peak from 300-400 ms.

The P300 paradigm is based on an ERP component that occurs as a positive deflection in voltage roughly 300 ms after a stimulus event [36]. A typical P300-based BCI design embeds an infrequent stimulus (the rare event) in a series of background stimuli (usual events). The method of recording the elicited response to rare events is referred to as oddball paradigm [28]. The P300 wave induced by a rare event is most prominent in the EEG recorded by electrodes covering the parietal area of the brain. By averaging over multiple recordings, the P300 amplitude is isolated from the noise. Figure 2.5 shows example signals recorded after a rare event and after a usual event.

Farwell and Donchin [28] developed the first P300-based BCI speller in 1988, where the oddball paradigm was implemented using visual stimulation in the form of flashing rows and columns (see Figure 2.4 B).

The P300 can be evoked in a high number of users: According to a study conducted by Guger et al. [37] in 2009, approximately 89% of the general population reach accuracies above 80%. As P300-based BCIs employ the flashes at a fixed pace, which the user cannot influence, they are considered as synchronous systems. Moreover, as these systems rely on external stimuli, they can be categorized as exogenous systems. With modern P300-based BCI applications, ITRs in the range from 12-70 bpm have been achieved [13, 38] (ITR is a standard performance metric for BCIs that integrates speed and accuracy into a single measure; for more details, please refer to section 2.7). Like MI-based systems, P300-based BCIs have been successfully tested with patients with advanced ALS [39].

### The VEP Paradigm

VEPs have been researched since the 1970s [11]. The VEP paradigm can be divided into several subgroups according to the stimulation modality. Among these subgroups are the SSVEP and the c-VEP paradigms, which are the main focus of this thesis.

SSVEPs, which are also called frequency-modulated visual evoked potentials (f-VEPs), are responses to a periodic visual stimulus occurring at the occipital and parietal cortical areas of the brain [40–42]. If the frequency of a flickering stimulus is high enough (> 6 Hz), the individual responses to each flash overlap; as a result, a steady-state brain response can be observed with dominant frequency components at the fundamental, harmonics and sub-harmonics of the stimulus frequency. Figure 2.6 A shows the amplitude spectrum for an SSVEP response to a stimulus flickering at 9 Hz as an example; peaks at 9, 18, 27, 36, and 45 Hz are clearly visible.

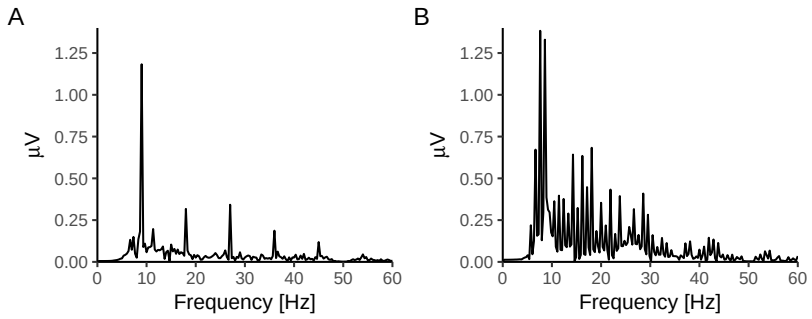



**Figure 2.6:** Amplitude spectrum for SSVEP and c-VEP response, calculated by fast Fourier transform (FFT). (A) The BCI user focused on a 9 Hz SSVEP stimulus. (B) The BCI user focused on a 63 bit $m$-sequence.

In BCI applications, the SSVEP paradigm is implemented in the following way. Multiple stimuli, each flickering with a unique frequency, are presented to the user while brain responses are recorded via EEG. By analyzing the brain signals, the system detects which target is gazed at by the user; it identifies the target frequency out of all stimuli as the one that is most dominant in the spectrum of the recorded signal. Each stimulus is linked to a control command specific to the application. In spellers, for example, multiple boxes with letters are used as stimuli; the desired letter can be selected by gazing at the corresponding box containing it. Figure 2.4 C shows the Bremen-BCI spelling application, which presents five selectable SSVEP targets for cursor navigation.

Regarding the SSVEP stimulation frequencies, several factors need to be considered: First, only a limited frequency range evokes a strong enough SSVEP response [43]; second, the harmonics of the frequencies might interfere with each other [40]; third, the generation of stable flickering patterns depends on hardware. For example, if computer screens are used for stimulus presentation, the frequencies depend on the refresh rate [6].

A similar VEP BCI approach is the c-VEP paradigm, which has initially been proposed by Sutter [44]. In contrast to SSVEP-BCIs, all stimuli are modulated with different time lags of the same binary code sequence. The states of the target objects, shown and not shown, correspond to the bits of the corresponding code. A particular pseudorandom binary sequence, a so-called m-sequence, is typically used because of its correlation property (an $m$-sequence is nearly orthogonal to shifted versions of itself, see section 3.2).

For classification, usually, a template matching approach is used [45]. When the user gazes at one of the stimuli, the system compares the recorded EEG data to templates, selects the best matching one, and produces the corresponding output. The EEG templates are generated a priori by averaging multiple trials collected in a training session during which the user needs to gaze several times at the stimuli. For the c-VEP paradigm, stimulus onset needs to be determined precisely. For this, data collection

and stimulus presentation need to be synchronized. Figure 2.6 B shows the amplitude spectrum for a typical c-VEP response.

Sutter [44] presented the first c-VEP-based BCI in 1984, which employed a microcomputer with MC68000 processor for signal analysis and a specially designed video display for stimulus presentation with up to 128 targets. This system has been tested by an ALS patient using implanted electrodes; the patient achieved communication rates of 10 to 12 words per minute with the BCI [7]. More than a decade later, in 2009, Bin et al. [45] implemented an EEG-based 16-target c-VEP system using a standard computer screen and reached a mean ITR of 92.8 bpm. By doubling the number of targets and using spatial filters based on canonical correlation analysis (CCA), two years later, Bin et al. [46] yielded a mean ITR of 108 bpm. Due to further improvements in classification and in the design of spatial filters, Spüler et al. [9] reached an even higher mean ITR of 144 bpm.

In comparison to the other BCI paradigms, VEP-based BCIs require little or no user training and achieve high selection speeds [41]. VEP-based BCIs are exogenous, as they require external stimuli. Moreover, almost all VEP-based BCIs belong to the category of dependent systems, as they typically rely on gaze direction controlled by extraocular muscles. Therefore, they may not be suitable for people with neuromuscular disabilities who cannot control the required muscles. Despite that, systems where VEPs were modulated by spatial attention only, independent of extraocular muscle control, have also been developed. An example is an SSVEP system proposed by Kelly et al. [47], which detects the desired target (one of two stimuli on the left and right side of the screen) only by means of visual spatial attention. In comparison to the more common dependent VEP BCIs, this system yields much lower ITRs and requires more user training.

While not addressed in this thesis, transient VEPs (t-VEPs), steady-state motion VEPs (SSMVEP), and motion-onset VEPs (m-VEPs) also belong to the category of VEP-based BCIs.

The t-VEPs are elicited by mutually independent flash sequences [45, 48]. Their classification involves averaging over multiple flashes. To prevent overlapping of t-VEPs, the stimulus rate of the flash sequence needs to be below 4 Hz. The achievable ITRs for t-VEP-based BCIs are around 30 bpm [45] (much lower than SSVEP and c-VEP-BCIs).

The m-VEPs are elicited by predefined motions of the visual targets and are usually comprised of three main peaks P1, N2 (predominantly motion-specific, latency 160-200 ms), and P2 (elicited by complex moving stimuli, latency 240 ms) [49].

The SSMVEPs are elicited when gazing at objects with sinusoidal-based movement patterns, e.g., contractions or oscillations. In SSMVEP-based BCIs, all stimulus objects are moving simultaneously with individual frequencies [50]. An advantage of SSMVEP-based BCIs over SSVEP-based BCIs is a more subtle stimulation; a disadvantage is the comparably low system speed.

## 2.4 General BCI Frame Work

Regardless of the neurophysiological principle used, the basic design of a BCI consists of several standard components. Figure 2.7 illustrates a general BCI framework, as introduced in [51].

In the following, a short description of the individual components is provided:

**User:**
The BCI user is the person who controls the system via brain signals, which either occur as a response to external stimuli (exogenous BCIs) or through self-regulated intrinsically produced EEG features (endogenous BCIs).

**Signal acquisition:**
The brain activity is recorded; for example, in EEG-based BCIs, signal electrodes are used, which are connected to an amplifier that amplifies and temporally filters the recorded electrical signals.

**Preprocessing:**
The signal is segmented, i.e., an appropriate classification window is determined. Furthermore, as the raw EEG is noisy, temporal filtering methods may be applied to reduce various bio-electrical artifacts. To limit the signal analysis to a specific frequency range, band-pass filters can be applied. Most VEP-based BCIs use stimuli above 6 Hz; low frequencies (e.g., below 2 Hz) can be filtered as they do not carry relevant information. Moreover, notch filters which reject signals in a specific frequency range are applied to remove the power line interference (in Europe 50 Hz). In VEP-BCI research, Chebyshev and Butterworth filters are commonly applied filter implementations (see, e.g., [8, 52]).

**Feature extraction:**
Informative and non-redundant values are extracted from the preprocessed data yielding a feature vector. Methods of dimensionality reduction, e.g., principal component analysis, may be applied.

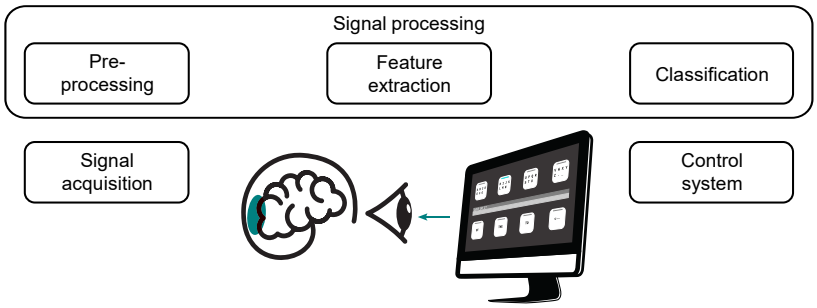



**Figure 2.7:** Components of a BCI. The general stages of a BCI system framework are shown.

**Classification:**

The signal features are translated into an output software command. This command is typically produced on the basis of established classification techniques such as Pearson's correlation method, linear discriminant analysis, support vector machines, and neural networks (see, e.g., [28, 53–55]).

**Control interface:**

The classifier output is translated into a control signal of the application. For example, in case of a spelling application, the character corresponding to the classified command is added to the screen; systems can additionally provide audio feedback.

It should be highlighted that the user is a critical component of the BCI. Various factors, such as age, tiredness, mental and physical health, can have an impact on the system performance. In respect to user variability, it is essential to customize the remaining components to suit individual needs. The investigation of user attributes and their impact on performance is relevant for the development of user-specific calibration methods.

## 2.5 BCI Applications

The earliest BCIs were mainly developed as communication tools for severely impaired patients [28, 44]. During the last decades, several other types of applications have been tested. Wolpaw et al. [12] categorized BCI applications as to whether their functioning is to replace, restore, enhance, supplement, or improve natural central nervous system (CNS) outputs. The following list provides examples to each of these categories:

**Replace:**

Spelling applications for people who lost control over their voice [13] and wheelchair control applications for people who lost control over limbs [5] belong to this category. The target group of these applications is patients who have

limited communication options (for example, late-stage ALS patients).

**Restore:**
Applications that electrically stimulate paralyzed muscles to restore limb movement or applications that stimulate peripheral nerves to restore bladder function belong to this category.

**Enhance:**
Non-medical applications that continuously monitor brain activity during demanding tasks, such as driving a car, belong to this category [56, 57]. These types of applications can be used to warn the user and enhance his or her attention in emergencies. The primary target group of this application scenario is healthy users.

**Supplement:**
In general, any application that provides an additional means of control belongs to this category. In computer games, where the user is controlling a joystick with both hands, the BCI can enable a supplementary selection function. BCIs could provide additional hands-free control mechanisms for augmented reality and virtual reality glasses. In smart homes, the artificial BCI outputs could be used to control lighting or entertainment devices [58, 59] while the user is manually engaged in other daily activities, e.g., preparing a meal. These application scenarios may be of interest to the entertainment sector.

**Improve:**
Rehabilitation applications with a focus on motor re-learning that detect and enhance brain signals from a damaged cortical area to stimulate muscles or improve movements belong to this category [60]. The target group includes stroke patients with impaired CNS functions. During motor exercises, the BCI can provide supportive feedback if the user performs the re-learned movement correctly.

The majority of current BCI research addresses the first type of BCI application, the replacement of lost CNS output. The thesis focuses on VEP-based spelling applications, which also belong to this category.

## 2.6 VEP-based BCI Spellers

According to the review of Rezeika et al. [13], only a fifth of the studies presenting BCI-spellers in the last decade directly addressed the design of the graphical user interface (GUI).

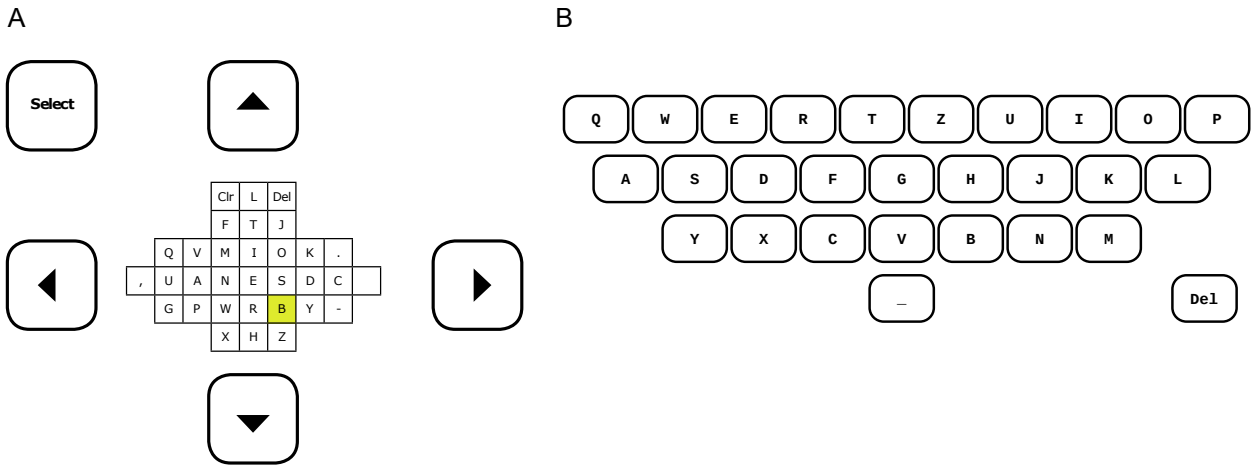



**Figure 2.8:** Multi-step speller and single-step speller. (A) The figure shows the Bremen speller [1]. Multiple cursor movements controlled via SSVEP are required to select a letter. (B) The figure shows the QWERTZ speller [61]. Only one SSVEP selection is required to select a letter.

The GUI design, however, is a crucial component in terms of BCI performance and usability. VEP-based BCI spellers can be categorized as to whether they are based on a multi-step or a single-step layout and as to whether they are synchronous or asynchronous. Additional GUI features, such as word completion or word suggestion modules, can improve the overall efficiency of the system.

### Multi-step Speller

The English alphabet consists of 26 letters. Typically, the number of stimuli used in SSVEP-based BCIs is much lower. For this reason, various multi-step interfaces, where the user needs to select several SSVEP targets to choose the desired character, have been developed. One example of a multi-step speller is the Bremen-BCI GUI, which presents a rhombus-shaped grid consisting of 32 characters [1]. Five SSVEP target stimuli (representing directions UP, DOWN, RIGHT, LEFT, and the command SELECT) allow the user to move a cursor along the grid and to select the desired character (see Figure 2.8 A). The usage frequency of characters in the English language determines their position in the grid; for example, E, the most frequent letter, is positioned in the center.

The Bremen speller was tested in a field study conducted at the RehaCare fair with 29 healthy participants and 8 participants with different kinds of disabilities [1]. In this study, an average ITR of 25.7 bpm was achieved. Over time, several modifications lead to an increased average ITR of 61.7 bpm with seven healthy participants and a peak ITR of 109 bpm [42].

Another example of a multi-step speller is the three-step speller we presented in [62, 63]. The three-step speller shows four boxes

as SSVEP stimuli to the user. One of these boxes contains the command DELETE, the other boxes contain selectable menus, each offering nine characters. For writing a letter, the user needs to produce three SSVEP commands. More details about this interface are provided in chapter 4.

The overall spelling speed of multi-step spellers is limited, as the classification windows and gaze-shifting periods of each step are accumulated.

### Single-step Speller

The earliest single-step c-VEP speller with multi-target stimulus presentation was developed in the 1990s by Sutter [7], who used a custom-designed cathode-ray tube (CRT) for stimulus presentation. In 2003, Gao et al. [40] presented a multi-target SSVEP system using 48 LED with a frequency resolution of 0.2 Hz to control a TV remote; their experiment demonstrated the feasibility of single-step SSVEP spellers. Hwang et al. [64] developed an SSVEP speller that resembled a QWERTY style keyboard. The authors used 30 LEDs flickering with different frequencies.

Meanwhile, researchers developed one-step spellers for standard liquid-crystal display (LCD) monitors; in 2011, Bin et al. [46] implemented a system with 32 targets based on the c-VEP paradigm.

While c-VEP stimuli share the same circular shifted code pattern, SSVEP stimuli need to differ in their cycle length; the update rate of the display hardware (i.e., the refresh rate) limits the number of suitable SSVEP stimuli. For this reason, SSVEP multi-target systems for standard monitors are harder to realize. In 2010, advancements in stimulus design led to the implementation of SSVEP one-step spellers for standard monitors. Wang et al. [65] developed a frequency approximation method to realize multi-target SSVEP systems. The authors tested the method with a 16-target virtual keypad yielding an ITR of 75 bpm with three participants. After further improvements in signal classification, Chen et al. [66] presented a one-step SSVEP speller, which yielded an average ITR of 267 bpm. In their system, 40 SSVEP stimuli were arranged as a $5 \times 8$ matrix allowing the selection of characters, numbers, and additional symbols.

While single-step spellers might be more intuitive and allow faster spelling speeds, they tend to be less accurate and cause more eye fatigue than systems with a low number of targets. In chapter 4, the impact of the number of targets on BCI performance, which is closely related to the number of steps of the GUI, is investigated.

### Synchronous Spellers

Synchronous VEP spellers produce system outputs at fixed time intervals. The flickering stops after a fixed time interval during which the system collects the EEG data. For example, for the c-VEP paradigm, this time interval is determined by the code-length of the $m$-sequence. After the flickering phase, the BCI produces an output by analyzing the collected data. Spelling applications typically incorporate a flickering pause after generating a command. During this phase, the flickering stops, and the user is given a fixed time interval to shift the gaze to the next letter. In the literature, this stimulation pause has been termed gaze shifting period [67], cue duration [68], break between trials [69], or rest period [47]. After the gaze shifting phase, the flickering continues. In this manner, gaze-shifting and flickering-phases alternate; the BCI generates outputs at equidistant time points which the user cannot influence.

Synchronous systems do not consider that a user – for any reason – is not looking at the target. Particularly, in spelling applications, the duration the user needs to locate the desired letter and to shift his or her gaze depends on many factors, such as familiarity and complexity of the letter arrangement, external distractions, and tiredness. Thus unintended selections might occur, an issue which is often referred to as the Midas touch problem (see, e.g., [67]). While synchronous system implementations lead to fast spelling speeds in short sessions, they might be impractical in long-term daily use.

### Asynchronous Spellers

Asynchronous VEP spellers can distinguish between intended target fixations and exploratory target fixations. These kinds of systems incorporate a no-control state [70] (also called idle state [71]), where the user does not intend to make a selection. With asynchronous applications, the BCI can provide continuous feedback reflecting the classifier state to the user [47].

For the SSVEP paradigm, asynchronous systems have been realized by introducing classification thresholds [42, 63, 69]. The system determines prediction scores after specific calculation intervals and compares these against threshold values. For example, in the Bremen-BCI, cursor movements or selections are only performed if the calculated frequency power estimation associated with the presumably fixated target surpasses a pre-defined threshold; otherwise, the system rejects the classification and collects further data [42]. The system also provides continuous feedback by varying the size of the targets in relation to the corresponding probability. In addition to the threshold criterion, pseudo-targets

can be implemented to improve the distinction between control and no-control state [71]. In this sense, the classifier calculates probability scores for the target frequencies displayed on the screen, and also, for frequencies that are not shown but might correspond to VEPs elicited when the user is scanning through the letters. If the BCI classifies any of these pseudo-targets, it does not generate an output. Typically, mean values between target frequencies are employed as pseudo-targets (see, e.g., [63, 70]).

For c-VEP-based BCIs, asynchronous systems are harder to realize; static time windows are the standard because of the required synchronization between EEG data collection and stimulus representation, the fixed length of the code sequence, and the dependence on pre-recorded data sets. In spite of these issues, we developed an asynchronous c-VEP speller [72], which is described in chapter 5.

In general, classification thresholds can lead to a more natural interaction between user and BCI, but they can slow down the output speed, as it takes extra time until they are surpassed. This performance drop may be compensated using word prediction methods.

### Spellers with word prediction features

Word completion and word prediction features allow users to produce outputs with fewer selections and can, therefore, speed up communication via BCI. So far, spellers offering predictions showed promising results. The majority of prediction methods have been developed for P300-based BCIs [73–75].

Ryan et al. [74] developed a P300-based spelling application that presented an $8 \times 9$ matrix for character, letter, and number selection. In addition to that, the GUI presented up to seven word-suggestions, preceded by a number, in a separate window. Users could choose the desired word by selecting the corresponding number in the matrix. According to their results, their word prediction mechanism led to an improved character output. However, the authors also observed reduced accuracy, which might be explained by an increased workload when using the additional prediction module. To reduce the cognitive load, Kaufmann et al. [75] developed a similar $6 \times 6$ P300 speller. In their system, the word suggestions were directly integrated into the matrix rather than in a side window. This modification led to a reduced workload on the user and yielded overall better performance.

For VEP-based BCIs, word prediction mechanisms are quite rare [73]. Regarding the SSVEP paradigm, Volosyak et al. [76] presented a dictionary functionality for the Bremen-BCI speller. The original Bremen-BCI speller was extended by an additional

drop-down list containing six dictionary suggestions, and a sixth SSVEP target box, which led to a menu layout. This menu layout presented each of the six suggested words in a flickering box; the user could select the desired word by focusing on the box containing it. For most users, the modified dictionary-driven Bremen-BCI realization led to an increased spelling performance.

Spelling interfaces based on the c-VEP paradigm are typically implemented as synchronous systems. Due to the additional time needed to check the dictionary suggestions and to locate the desired target, dictionary integrations are more suitable for asynchronous spellers such as the Bremen-BCI. The developed asynchronous c-VEP system [72] employs word suggestions on the word level. The functionality of the dictionary feature is also described in detail in chapter 5.

## 2.7 Evaluation Metrics

Various evaluation metrics can be applied to analyze BCI performance. Among the most common evaluation metrics for BCI spelling applications are the classification accuracy, the ITR, and the OCM. Each of these measures has its own advantages and disadvantages [77].

### Classification Accuracy

To determine the accuracy of a BCI classifier, one can investigate the confusion matrix (see, e.g., [77]). For two-class problems, the confusion matrix partitions the classifications into true positive (TP), false positive (FP), false negative (FN), and true negative (TN) recognitions (see Figure 2.9 A). Various performance metrics can be derived from this matrix. Most commonly used in BCI research is the classification accuracy, $P$, which describes the probability of

A

| Predicted Values \ Actual Values | Positive (1) | Negative (0) |
|---|---|---|
| Positive (1) | TP | FP |
| Negative (0) | FN | TN |

B

| Class | 1 | 2 | 3 | 4 | 5 | Total |
|---|---|---|---|---|---|---|
| 1 | **17** | 1 | 0 | 1 | 1 | 20 |
| 2 | 2 | **15** | 1 | 0 | 2 | 20 |
| 3 | 0 | 1 | **18** | 1 | 0 | 20 |
| 4 | 0 | 0 | 0 | **20** | 0 | 20 |
| 5 | 1 | 2 | 1 | 1 | **15** | 20 |

**Figure 2.9:** (A) General confusion matrix for a two class problem ($N = 2$). The observations are categorized into true positives (TP), false negatives (FN), false positives (FP), and true negative (TN). (B) Example of a confusion matrix for multiple classes (here, $N = 5$). In the example, an accuracy of 85% is achieved (calculated as the sum of diagonal elements divided by the total sum of cases).

a correct classification; in the binary case,

$$P = \frac{TP + TN}{TP + FP + FN + TN}. \tag{2.1}$$

More generally, for $N$ classes, the confusion matrix $(n_{ij})$, with rows and columns $i, j = 1, \ldots, N$, displays the relationship between the actual class the user intended to select, and the predicted class, determined by the classifier. A single element $n_{ij}$ indicates the number of occurrences the actual class $i$ was classified as class $j$. Hence, the diagonal elements, $n_{ii}$, represent the number of correct classifications for the class $i$. The accuracy can, therefore, be calculated as the sum of diagonal elements divided by the total sum of elements,

$$P = \frac{\sum_{i=1}^{N} n_{ii}}{\sum_{i=1}^{N} \sum_{j=1}^{N} n_{ij}}. \tag{2.2}$$

An example confusion matrix for a 5-class system (such as the SSVEP Bremen speller [42]) is depicted in Figure 2.9 B. As equation (2.2) reveals, the accuracy depends on the number of classes and the individual frequency of cases. The chance level, i.e., the threshold that denotes if the accuracy is better than random is depended on the same values. Ideally, data should be balanced in the sense that each actual class has the same number of occurrences. The following example illustrates why unbalanced data sets are not desired. We consider a binary classification problem with 10 cases where the actual class is 0 and 30 cases where the actual class is 1. The classifier that only outputs class 1, independently of the actual class, achieves 75% accuracy; the classifier that only outputs class 0, independently of the actual class, achieves 25% accuracy. Due to the imbalance of the data, a comparison of the classifiers based on the accuracy is biased.

### Cross-Validation

Cross-validation (also called rotation estimation) describes a set of validation techniques for classification models that were trained with a dataset of known labels [78]. They are typically used to estimate a statistical measure (such as the accuracy) of a classification model on unknown data (i.e., data that was not used for classifier training).

In cross-validation, a sample of data is partitioned into subsets. Usually, several rounds of cross-validation are implemented using different partitions. The results are averaged across these rounds.

In $k$-fold cross-validation, a dataset of observations of known labels is partitioned into $k$ equal-sized subsets (i.e., each subset contains

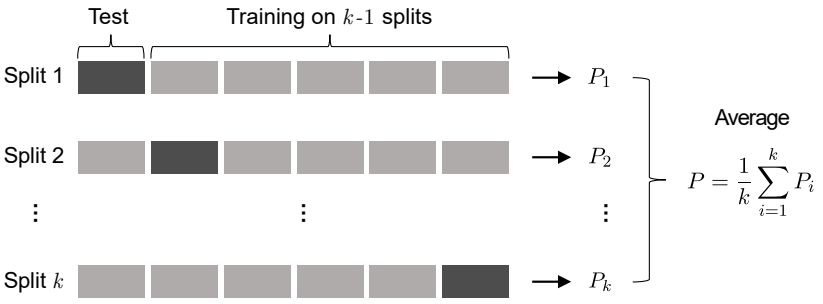



**Figure 2.10:** $k$-fold cross-validation. In $k$-fold cross-validation, the data are divided into $k$ complementary subsets. The $i$-th set is used as test set $i = 1, \ldots, k$. The remaining $k - 1$ subsets are used as training sets. For each round (i.e., each of the $k$ sets used as test set), a performance metric (here, $P_i$) is calculated. At last, a mean value across all rounds is calculated.

the same number of observations). From these, $k-1$ subsets (called the training folds) are used to train the model. The subset that is left (called the validation fold or test fold) is used to evaluate the model of the classifier. In total, $k$ of such validation rounds are performed. Each round a different subset is used as test data, and the remaining $k - 1$ subsets are used as training data. An overall average across the results of the different rounds is calculated, yielding a single estimation. Note that all observations are used for training and validation. Figure 2.10 illustrates the procedure of $k$-fold cross-validation.

In the case where $k$ is the number of observations, the method is called leave-one-out cross-validation. In stratified $k$-fold cross-validation, in each fold, the labels of the observations occur equally frequent.

There are two use cases for cross-validation. On the one hand, it can be used to compare two different prediction models. For example, to compare two classification methods for a BCI spelling application, for each of these methods, cross-validation can be performed using labeled off-line data to estimate the classification accuracy. On the other hand, cross-validation can be used to optimize a parameter specific to a prediction model. In this sense, cross-validation is performed multiple times with the same model but with different values for the parameter. The parameter can then be set to the value that maximizes the averaged results from the cross-validation (e.g., the parameter value yielding highest accuracy).

### Information Transfer Rate

The ITR is a standard metric to examine BCI performance. Its calculation is based on Shannon's channel theory for general communication systems [79] and has been introduced by Wolpaw et al. [12]. The ITR is a measure of the mutual information between the user's choice and the BCI selection; as such, it interprets the BCI system as a noisy channel, where noise is added whenever a misclassification occurs [80].

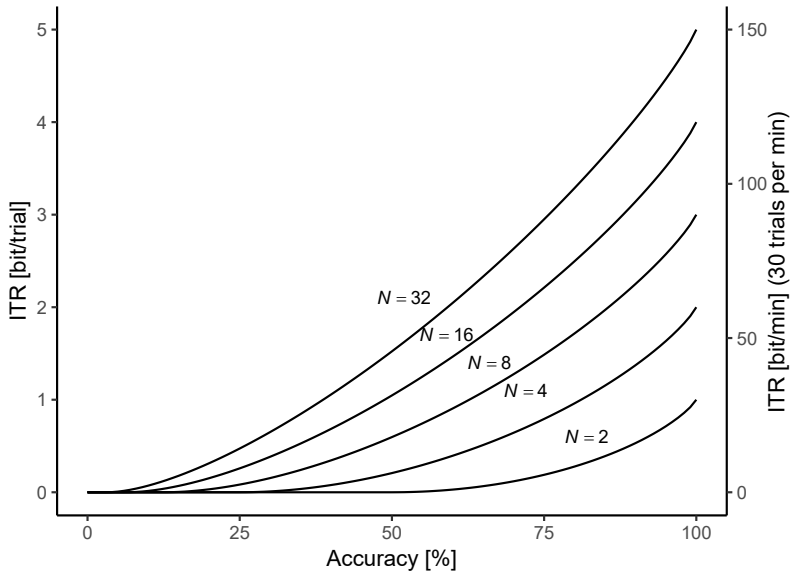



**Figure 2.11:** Information transfer rate in bit/trial for different numbers of classes, $N = 2, 4, 16, 32$. Provided are also the associated values for the ITR in bit/min (assuming 30 trials/min, i.e., a selection time of 2 s). The bit rates are only shown for accuracies greater than or equal to chance level (i.e., $P \geq 1/N$). Figure adapted from Wolpaw et al. [12], Figure 1.

The measure combines information on system speed and accuracy in one variable while also taking into account the number of classes. The metric is expressed as the number of error-free bits per time unit. The ITR in bit/min, $B_m$, can be calculated using the following formulas:

$$B_t = \log_2 N + P \log_2 P + (1 - P) \log_2 \left( \frac{1 - P}{N - 1} \right), \tag{2.3}$$

$$B_m = \frac{60}{T} \cdot C \cdot B_t, \tag{2.4}$$

where $B_t$ denotes the information transferred in bit per trial, $N$ the number of classes, $P$ the classification accuracy, $T$ the total time of the experiment task, and $C$ the number of classifications of the experiment task.
A calculator for the ITR can be found under the tools section of `https://bci-lab.hochschule-rhein-waal.de/en/itr.html`. In this thesis, the unit of the ITR, bit/min, is abbreviated as bpm.

Figure 2.11 shows the ITR in bit per trial as a function of the classification accuracy for different numbers of classes. The figure also shows the associated values of the ITR in bit per minute if the average classification time is 2 s per trial (i.e., 30 trials/min). It can be seen that the ITR increases with the number of classes $N$. In the following, the formula (2.4) is examined in detail. The first summand of the formula, $\log_2 N$, is strictly positive and independent of the classification accuracy $P$. The second summand, $P \log_2 P$, contributes negatively to the ITR and is independent of the number of classes. The third summand of the ITR formula, $(1-P) \log_2 \left( \frac{1-P}{N-1} \right)$, is dependent on the classification accuracy $P$ and on the number of classes $N$. The term also contributes negatively to

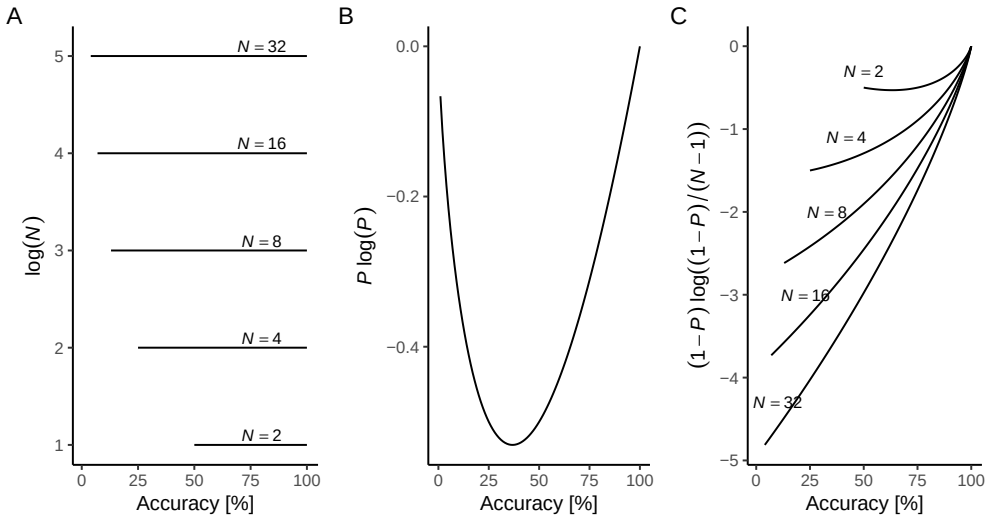



**Figure 2.12:** Individual terms of the ITR formula. (A) The first summand, which depends only on the number of classes. (B) The second summand, which depends only on classification accuracy. (C) The third summand, which depends on both the classification accuracy and the number of classes. Values are only shown for accuracies greater than or equal to chance level (i.e., $P \geq 1/N$). Figure adapted from `https://bci-lab.hochschule-rhein-waal.de`.

the ITR for accuracies below 100%. However, for fixed accuracies, the values increase with the number of classes $N$. Figure 2.12 shows the individual summands as a function of the classification accuracy for different numbers of classes.

For multi-step spellers, the number of classes can be either set as the number of frequencies (i.e., the number of choices on the level of selections in each step) or as the total number of possible selections (i.e., on the level of output characters) [1, 38, 70]. With the first option, the ITR describes the performance of the raw BCI and reflects the efficiency of the classification algorithms; with the second option, the ITR reflects the effectiveness of the application.

Strictly speaking, the ITR calculation is done under the assumptions that all possible selections are equally probable and that the system is memoryless [12, 80, 81]. As practical experiments often violate these assumptions, for example, when testing systems with word prediction models, researchers usually provide additional performance metrics.

### Output Characters per Minute

The ITR alone is not an appropriate measure for the evaluation of the application usability. In systems where many misclassifications occur, the ITR values may be too high, as the following example demonstrates: We consider a P300 speller such as the Farwell and Donchin speller [28] with $N = 36$ selection options, which is tested with two setups. In the first setup, the selection time for each letter

is 1 s. With these parameters, the user achieves an accuracy of 30% after 100 selections. In the second setup, the selection time for each letter is 10 s. Now the user achieves an accuracy of 100% after 100 selections. Calculation of the ITR results in 41.9 bpm for the first and 31.0 bpm for the second setup. Although the first setup is not usable with a classification accuracy as low as 30%, it yields a higher ITR.

In comparison to the ITR, the measures introduced in the following reflect the application speed in terms of output characters rather than in terms of selections.

Correct letters per minute (CLM) is a measure for single-step spellers that determines the theoretical number of correct letter selections per minute by simulating the correction of typing errors [25, 81]. In single-step spellers, where an erroneous selection requires one additional correct choice (delete the previous character), the number of correct characters can be calculated as the difference of correct selections and incorrect selections. Taking the classification accuracy $P$ into account, the number of characters is $PC - (1 - P)C = (2P - 1)C$, where $C$ refers to the total number of selections. If the accuracy is below 50%, the number of errors is higher than the number of correct selections; the system is not usable. Thus the CLM metric is defined as

$$CLM(T, P) = \begin{cases} \frac{60(2P-1)}{T} & P > 0.5 \\ 0 & P \leq 0.5 \end{cases} \tag{2.5}$$

where $T$ refers to the average classification time for one selection in seconds, and $P$ refers to the classification accuracy. The unit of the CLM is characters per minute, abbreviated char/min.

Back to the previous example, if we calculate the CLM, the first setup ($T = 1$, $P = 0.3$) results in a CLM of 0 char/min and the second setup ($T = 10$, $P = 1$) results in a CLM of 6 char/min. Therefore, in this example, the CLM reflects the usability of the application better than the ITR.

The CLM metric is not suitable for multi-step spellers where the number of selections for error corrections depends on the current step or the position of a cursor. Moreover, the CLM can only be used if one selection corresponds to exactly one character that is added to the output display. It is, therefore, not suitable for interfaces where selection options can represent entire words.

For this purpose, Ryan et al. [74] suggested a similar measure, the output characters per minute (OCM), which is simply calculated by dividing the total number of output characters by the time required to spell them. This metric can be applied to measure the

efficacy of multi-step spellers and dictionary-supported spellers, where a single selection can represent an entire word. To determine the OCM, the user needs to correct all errors. The OCM is also expressed in characters per minute (char/min). It is directly dependent on selection accuracy, and, similarly to the CLM, it provides a better measure in terms of system output in comparison to the ITR.

The information value of the different measures depends on what aspect of performance is investigated. To assess the performance of the classifier, the ITR is a suitable measure. To assess the performance of the spelling application, OCM and CPM may be the better option. In this thesis, we mainly use ITR (on the level of the selections in each step), the classification accuracy, and the OCM. These three measures cover the performance on the classification level and the application level.

## 2.8 BCI Illiteracy

BCIs not always interpret the user's intent with sufficient accuracy. Spelling applications are not usable if the average classification accuracy is too low. The phenomena where no sufficient control over the system is achieved has been termed BCI illiteracy [8, 29, 82].

A BCI user is referred to as BCI-illiterate if the system fails to detect his or her intentions accurately, more precisely, if the classification accuracy does not surpass a certain threshold.

It should be noted that this threshold value is neither standardized nor fully justified in BCI research [83]. A threshold for reliable control is indeed difficult to justify, as it depends on the number of targets, but also on the system design and system purpose. This thesis uses a threshold of 70% accuracy to define BCI-illiteracy, a value that has often been employed in this regard in the literature [84, 85].

The BCI literacy rate is defined as the percentage of users who achieve control over the system, and the BCI illiteracy rate is defined analogously [86]. It is estimated that across all major BCI approaches the BCI illiteracy rate is around 15-30% [29].

The term BCI illiteracy is criticized for implying that it is the user's fault that he or she cannot control the system. However, BCI illiteracy is not an issue of the user but of the BCI that is not able to interpret the user's intent. Some researchers prefer other descriptions of the illiteracy phenomena instead, e.g., "lack of BCI

efficiency" [87]. The categorization into BCI illiterate and BCI literate users is useful to identify reasons for insufficient performance. Efforts to solve the BCI illiteracy problem include improvements of signal classification algorithms, interface design, training protocols (clear instructions, extended training of user and classifier), and sensors [88]. Addressing these points improves the system accuracy and thus benefits all users. This thesis mainly focuses on the first point (classification algorithms) but also addresses interface design and training protocols.

Often, algorithms are tested with existing datasets using off-line evaluation methods such as cross-validation [83]. While this approach is an effective way to evaluate and compare different classification methods, on-line tests are essential to validate overall BCI performance, as it takes instability over time, distractions due to the real-time feedback as well as environmental disturbances that affect performance in practical scenarios into account. The evaluation of BCIs should, therefore, incorporate an on-line session where the BCI is tested under real-world conditions [38].

While BCI literacy is defined on the basis of accuracy, Volosyak et al. [1] introduced the term BCI ability rate as a function of the ITR measure, representing the percentage of the population that can achieve a particular desired ITR value. Similar to the BCI literacy rate, the BCI ability rate can be seen as characteristic of the specific tested BCI system.

# Implementation of a VEP-based BCI

# 3

In this chapter, the typical implementation of VEP-based BCI systems for a typical setup consisting of a computer, signal amplifier, and a standard computer monitor is described. In this regard, the methods used in publications [72, 86, 89] are summarized. Particular emphasis is put on stimulus presentation and the classification of the EEG data, which are running in separated dedicated threads (i.e., independently running functions executed in parallel).

Multichannel methods such as the minimum energy combination (MEC) and the CCA are among the most widely used BCI signal processing methods and have been tested in many practical experiments. Both methods use linear combinations of individual electrode channels to generate one or more filtered channels. Another critical aspect of the implementation of a VEP-based BCI is the stimulus pattern.

The chapter presents different methods of stimulus presentation for SSVEP-based BCIs, such as frequency approximation and sinusoidal modulation techniques (section 3.1). For the c-VEP stimulus presentation, the generation and properties of the underlying $m$-sequence are discussed in detail (section 3.2). After that, the chapter focuses on the classification methods that were used in this thesis. The algorithms for the MEC (section 3.3) and CCA (section 3.4) are provided. Thereafter, a template matching method, which can be used for both c-VEP and SSVEP-based BCIs, is described (section 3.5). Following that, the chapter covers the synchronization between stimulus presentation and data acquisition (section 3.6). The last section discusses different training aspects of the classification methods applied in VEP-BCI research (section 3.7).

## 3.1 Presentation of Visual Stimuli for SSVEP-based BCIs

In SSVEP-BCIs, each selectable target flickers at a specific frequency. To generate the flickering pattern on conventional monitors, researchers used various methods; the most common approaches are presented in this section. An important component for the stimuli presentation is the vertical refresh rate of the monitor. The vertical refresh rate is the number of times per second the monitor draws images using the data it is given. The higher the refresh rate, the smoother the transition of images. Typical computer monitors have

a 60 Hz refresh rate; modern gaming monitors have refresh rates of 120, 144, or 240 Hz.

### Divisors of the Vertical Refresh Rate

The stimulus pattern can be obtained from integer divisors of the vertical refresh rate, which define the period lengths of a possible frequency [42, 63, 90]. Let $r$ denote the monitor refresh rate in Hz, and $D$ denote an integer divisor of $r$. The stimulus sequence with period $D$ representing the frequency $f = r/D$ is given by

$$c(i) = \begin{cases} 1 & i - 1 \pmod{D} < \left\lceil \frac{D}{2} \right\rceil \\ 0 & \text{otherwise,} \end{cases} \quad i = 1, 2, \ldots, \tag{3.1}$$

where $\lceil x \rceil$ denotes the ceiling function, which maps $x$ to the least integer equal to or greater than $x$.

The render elements, which are presented on the monitor, are determined within the software thread dedicated to the graphical presentation. The iterator $i$ of the stimulus sequence is updated every frame. If $c(i) = 1$, the stimulus is drawn to the screen; if $c(i) = 0$, the stimulus is not drawn.

Some research groups found that the duty cycle (i.e., the fraction of an 'on-off' period where the stimulus is 'on') might influence BCI performance (see, e.g., [91]). When employing formula (3.1) for stimulus presentation, the duty cycle is given by $\lceil D/2 \rceil / D$. Hence, for odd $D$, the duty cycle of the sequence is not 50%. For example, if the monitor refresh rate $r$ is 60 Hz, a cycle length of $D = 3$ yields a 20 Hz stimulus pattern with a duty cycle of 2/3=66%.

**Table 3.1:** Suitable SSVEP frequencies for 60 Hz displays. The frequency is calculated as $60/D$, where $D$ denotes the period length in frames.

| Period length | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|
| Frequency [Hz] | 20 | 15 | 12 | 10 | 8.57 | 7.5 | 6.67 | 6 |

The overall number of distinct stimuli that can be generated with divisors of the vertical refresh rate is limited. As shown in Table 3.1, with a vertical refresh rate of 60 Hz, only eight frequencies between 6 and 20 Hz, a typical range for SSVEP stimuli choice (see section 4.1 for more details), can be realized. Nonetheless, when using a 120 Hz refresh rate, the number of realizable frequencies in this range increases to 15.

Additional limitations can arise if the classification algorithms consider harmonics of a stimulus frequency. For example, 6 Hz and 12 Hz stimuli should not be used simultaneously. More generally, the following restriction rules need to be considered when choosing

SSVEP stimulation frequencies to avoid mutual influences (see, e.g., [92]):

$$f_i \neq 2f_j, \quad f_i \neq (f_j + f_k)/2, \quad f_i \neq 2f_j - f_k. \tag{3.2}$$

Due to these limitations, the method is not suitable for multi-target BCI applications.

### Frequency Approximation Method

A method to generate arbitrary frequencies up to half of the monitor refresh rate was first described by Wang et al. [65]. The authors implemented a 16-target system with a frequency resolution of 0.25 Hz and achieved an average ITR of 75.4 bpm.

Their so-called frequency approximation method employs a square wave function with amplitudes alternating at frequency $f$, which is given by

$$s_f(t) = 2(s\lfloor ft \rfloor - \lfloor 2ft \rfloor) + 1, \tag{3.3}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. The code sequence that is used for the stimulus presentation of the frequency $f$ is now given by

$$c(i) = \frac{1}{2}\left(1 + s_f\left(\frac{i-1}{r}\right)\right) \quad i = 1, 2, \ldots, L, \tag{3.4}$$

where the code length $L$ is a multiple of the frame rate $r$, inversely dependent on the frequency resolution [66]. For example, with a frequency resolution of 0.5 Hz (e.g., 6 Hz, 6.5 Hz, 7 Hz, and so on) the code length is set to $L = 2r$. Figure 3.1 shows the stimulation cycles of a generated 7, 9, and 11 Hz stimulus.

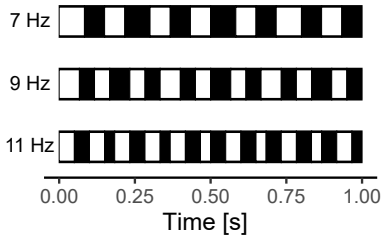



**Figure 3.1:** Example of stimuli generated using the frequency approximation method. Displayed are the 1 s stimulation cycles of a 7, 9, and 11 Hz stimulus for a refresh rate of 60 Hz.

A characteristic of this method is that the number of frames that represent a period of an approximated frequency varies. As a result, the duty cycle is not constant, as can be seen, for example, when inspecting the code sequence for the approximation of a 7 Hz stimulus with refresh rate 60 Hz:

$$\underbrace{111110000}_{9}\,\underbrace{111100000}_{9}\,\underbrace{11110000}_{8}\,\underbrace{111100000}_{9}\,\underbrace{11110000}_{8}\,\underbrace{111110000}_{9}\,\underbrace{11110000}_{8}.$$

The sequence consists of a total of seven 'on' and 'off'-phases of varying lengths, three consisting of eight frames and four consisting of nine frames.

One advantage of this method is that any frequency between 0 and $r/2$ Hz can be generated. As SSVEPs evoked by approximated frequencies with a low resolution of 0.1 Hz can still be reliably distinguished [40, 93], this method is suitable for the implementation of multi-target systems.

### Sinusoidal Stimulus Modulation

The manipulation of the transparency of a stimulus object can yield a more subtle flickering pattern. Manyakov et al. [94] proposed a sinusoidal stimulus modulation method, which is based on sampled sinusoidal intensity profiles. Using this technique, the stimulus sequence $c$ for the frequency $f$ is given by

$$c(i) = \frac{1}{2}\left(1 + \sin\left(2\pi f \frac{i-1}{r}\right)\right), \quad i = 1, 2, \ldots, L, \tag{3.5}$$

where the code length $L$ is a multiple of the frame rate $r$, as in the previous section. The values of this sequence range from 0 to 1.

The function values are used to manipulate the pixels corresponding to the stimulus. In this regard, alpha compositing, a technique used in computer graphics where different images are superimposed by manipulating the so-called $\alpha$ value, is applied. The color information of one pixel, usually represented by the RGB color values (red, green, blue), is extended to a fourth code that represents this $\alpha$ value [95]. This $\alpha$ value is a measure of transparency or opacity; it indicates how much a graphical element covers the elements on deeper levels. By manipulating the alpha value, semi-transparent images can be realized. The sum of transparency and opacity is always 1 (e.g., if the opacity of a graphical element is 60%, the transparency is 40%). The fineness of the transparency gradation depends on the number of bits used to store the alpha channel. For the most common image formats, one additional byte per pixel is used. In this case, the alpha channel comprises $2^8 = 256$ gradations, with the extreme values '0' for 'fully-transparent' and '255' for 'fully-opaque'.

For the implementation of an SSVEP stimulus $f$ the $\alpha$-values of the graphical stimulus element are set to $\alpha = [(c * 255)]$, where $c$ is a value of the code sequence obtained by (3.5) and $[\cdot]$ denotes the rounding operator, which rounds to the nearest integer (when the fraction part is precisely 0.5 it rounds up). Typically, for maximum contrast, the graphical stimulus elements are colored white, and the background is colored black. Thus, if $c = 0$, the stimulus is fully transparent, i.e., only the background (black) is visible at its place. On the other hand, if $c = 1$, the stimulus is fully opaque, i.e., it is colored white (maximal contrast to the background). For values between 0 and 1, the stimulus color goes through different shades of gray. The step width of the grayscale gradient is dependent on the monitor refresh rate (see Figure 3.2). At high refresh rates, a finer gradation is achieved, leading to a more subtle visual stimulation.

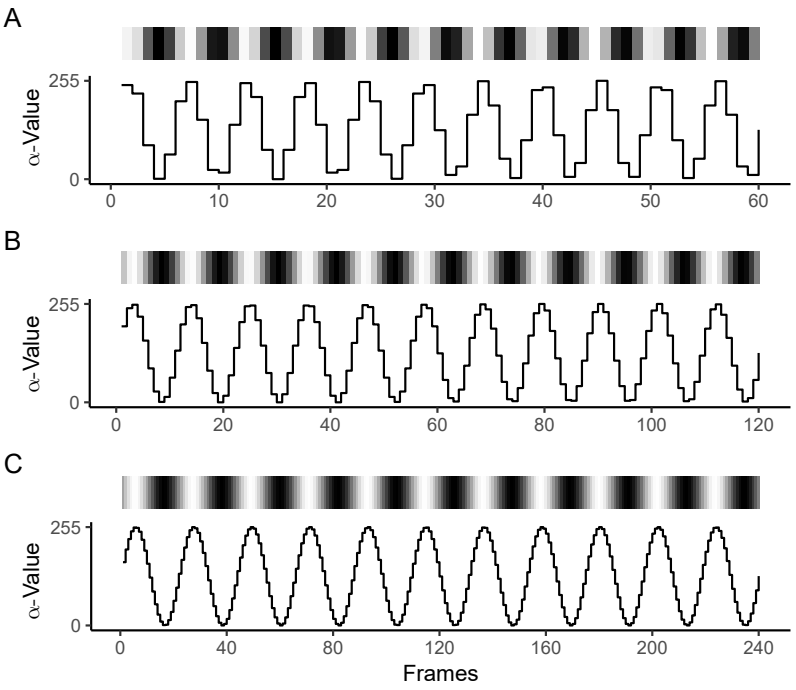



**Figure 3.2:** The $\alpha$-channel of an 11 Hz stimulus. Displayed are the $\alpha$-values during a one second stimulus interval at different monitor refresh rates. The alpha channel has 256 discrete levels , '0' denotes 'fully-transparent' and '255' denotes 'fully-opaque'. (A) Refresh rate 60 Hz. (B) Refresh rate 120 Hz. (C) Refresh rate 240 Hz.

### Hybrid Frequency and Phase Coding

The methods for SSVEP stimulus presentation described so far rely on frequencies for information coding. Phase information can also be used to code SSVEP targets. For example, Manyakov et al. [94] implemented a BCI with multiple targets at the same frequency but with different initial phases.

In the field of SSVEP research, the best results were achieved, when frequency and phase coding were combined [52, 66, 96]. In this sense, the approximation approach and the sinusoidal modulation method can be extended to the phase domain [96, 97].

When using the so-called hybrid frequency and phase coding approach [96] in combination with the sinusoidal modulation method, the stimulus sequence $c$ of frequency $f$ with phase $\Phi$ is given by,

$$c(i) = \frac{1}{2}\left(1 + \sin\left(2\pi f \frac{i-1}{r} + \Phi\right)\right), \quad i = 1, 2, \ldots, L, \qquad (3.6)$$

where, as before, $L$ and $r$ denote code length and refresh rate.

## 3.2 Presentation of Visual Stimuli for c-VEP-based BCIs

In c-VEP systems, all targets are modulated with cyclically shifted versions of a single code [45]. Maximum length sequences (*m*-

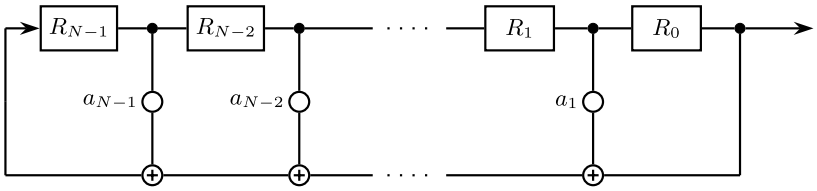



**Figure 3.3:** Basic $N$-stage linear feedback shift register (LFSR). The register cells hold binary states 1 and 0. The outputs are connected by XOR gates. Figure adapted from Gembler et al. [98].

sequences) are special non-periodic binary code patterns that are particularly suitable for the stimulus design in BCI applications because of their correlation property [7]. In the following, the generation of $m$-sequences and some of their properties are described.

### Generation of $m$-Sequences

A maximal-length sequence ($m$-sequence) is a periodic binary sequence with a noise like wave-form [99]. An $m$-sequence can be generated using a linear-feedback shift register (LFSR), as shown in Figure 3.3. An LFSR consists of $N$ binary memory stages (also called cells) labeled $R_{N-1}, \ldots, R_1, R_0$. The input bit, $R_{N-1}$, is the value of a linear function $f$ that performs modulo-2 additions with a subset of the register entries. (The modulo-2 sum of two bits is 0 if the bits are identical, and 1 if they differ, $0 + 0 = 0$, $0 + 1 = 1$ and $1 + 1 = 0$.) The register outputs of the LFSR are connected by exclusive-OR (XOR) gates. The bit positions that influence the next state (weights $a_i \neq 0$) are called taps.

A timing clock controls the memory stages of the LFSR. At each pulse of the clock, the states of the stages are shifted to the next stage. The entry in cell $R_i$ is passed to cell $R_{i-1}$, $i = N - 1, \ldots, 1$. The entry in stage $R_0$ (the rightmost register) determines the output of the LFSR. The sequence of output bits, $(b_i)$, is called the output stream. The entry in the leftmost register $R_{N-1}$ is updated with the value of the feedback function $f$. An LFSR must be initialized with a nonzero value, which is called seed.

A binary code of length $N$ can assume $2^N$ values. However, the period of the code produced by the LFSR can have a maximal length of at most $2^N - 1$, i.e., the LFSR cycles through all states except for the case where all bits are zero (in that case, the register states do not change). In the case where the period $L$ of the output stream has maximal length, $L = 2^N - 1$, the initial cycle of the generated output stream, $(b_0, b_1, b_2, \ldots, b_{2^N-2})$ is called maximal length sequence or $m$-sequence.

The combination of the register pins can also be expressed as mod 2 polynomial:

$$G(X) = X^N + a_{N-1}X^{N-1} + \ldots + a_2X^2 + a_1X + 1, \qquad (3.7)$$

| Timer | $R_2$ | $R_1$ | $R_0$ |
|---|---|---|---|
| 0 | 1 | 0 | **0** |
| 1 | 1 | 1 | **0** |
| 2 | 1 | 1 | **1** |
| 3 | 0 | 1 | **1** |
| 4 | 1 | 0 | **1** |
| 5 | 0 | 1 | **0** |
| 6 | 0 | 0 | **1** |
| 7 | 1 | 0 | 0 |
| ⋮ | ⋮ | ⋮ | ⋮ |

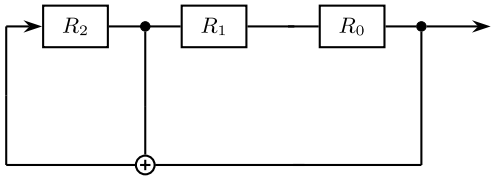



**Figure 3.4:** Generator polynomial and corresponding $m$-sequence. Shown is an $m$-sequence, that is obtained with the generator polynomial $g(x) = 1 + X^2 + X^3$. The seed (initial values of the registers) was set to 100. The generated $m$-sequence is 0011101.

where the coefficients $a_i \in \{0, 1\}$, $i = 1, \ldots, N-1$ denote the weights of the corresponding register pins. The length of the generated sequence depends on the values of the coefficients $a_i$. The LFSR produces an $m$-sequence ($L = 2^N - 1$), if and only if the corresponding generator polynomial is primitive, i.e., it can not be reduced to a product of polynomials of lower order and the smallest integer $k$ for which it divides $x^k - 1$ is $k = L$ (e.g., [100]). Figure 3.4 shows the generation of an $m$-sequence with the primary polynomial $g(x) = 1 + X^2 + X^3$ as an example.

### Properties of $m$-Sequences

The $m$-sequences have several desirable mathematical properties (see, e.g., [100]), some of which will be provided in the following. To describe these properties, we transform the sequence to a list of 1s and -1s, representing 'positive contrast' and 'negative contrast', respectively. More specifically, binary 0 is mapped to 1, and binary 1 is mapped to -1:

$$s_i = -2b_i + 1. \tag{3.8}$$

In the following, we consider the infinitely long periodic output stream $(s_0, s_1, s_2, \ldots)$ with a period of length $L = 2^N - 1$ generated by an $N$-stage LFSR.

**Balance property** The number of 1s (binary 0s) and -1s (binary 1s) only differ by 1. For an $m$-sequence of length $2^N - 1$ the number of 1s is $2^{N-1} - 1$ and the number -1s is $2^{N-1}$,

$$\sum_{i=0}^{2^N-2} s_i = -1. \tag{3.9}$$

**Correlation property** An $m$-sequence is almost completely uncorrelated with itself (i.e., nearly orthogonal to itself) for all time shifts:

$$\frac{1}{L} \sum_{i=0}^{2^N-2} s_i s_{i+j} = \begin{cases} 1 & j = 0, \\ -1/L & j \neq 0. \end{cases} \tag{3.10}$$

For arbitrary large code length $L$, the autocorrelation function approximates a Kronecker delta function, which is 1 if $i = j$, and 0 otherwise.

**Run property** A run denotes a string (or tuple) of consecutive 1s or a string of consecutive -1s. In an $m$-sequence, one-half of the runs have length 1, one-quarter have length 2, one-eighth have length 3, etc. (as long as these fractions give integral numbers). In each case, the number of runs of -1s is equal to the number of 1s.

The autocorrelation property is particularly advantageous for the implementation of VEP-based BCIs.

### Using $m$-Sequences in VEP-based BCIs

For VEP-BCIs, the brain responses evoked by different stimuli should be uncorrelated to each other, as this would enhance target discrimination. Because of the correlation property, patterns based on an $m$-sequence and its different time lags are a logical choice for BCI stimulus design.

Figure 3.5 shows the auto-correlation of an $m$-sequence and that of a typical averaged evoked response. It can be seen that the orthogonality of the stimuli patterns is not fully transferable to the corresponding VEPs. In practice, the responses from consecutive frames interact with each other [7]. Another issue is that the immediate neighbors to a fixated target also contribute to the response. The stimuli of a c-VEP-based BCI are usually arranged as a matrix (see Figure 3.6). In the case where other neighboring targets do not surround a target stimulus, i.e., at the boundaries, the evoked response might differ significantly from the response of the center targets. Therefore, some research groups implemented additional stimuli, which are not selectable, around the matrix of the target stimuli (Figure 3.6). This method is referred to as wrap-around principle or principle of equivalent neighbors [7, 101]. The idea behind this approach is to increase the similarity between the evoked responses when gazing at different targets.

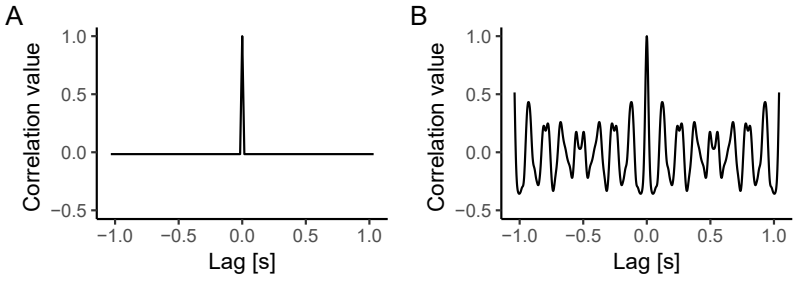
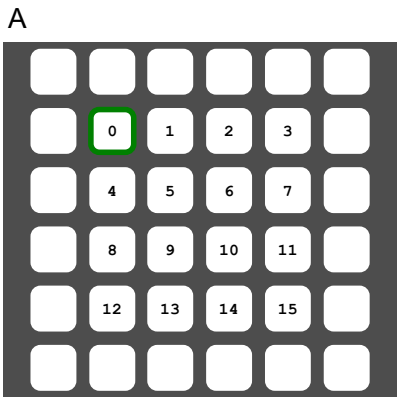



**Figure 3.5:** Autocorrelation property of the $m$-sequence and the evoked c-VEP. (A) Shown is the auto-correlation of the code pattern of a 63 bit $m$-sequence. (B) Shown is an example of the auto-correlation of the averaged evoked response.




B

| | | | | | |
|---|---|---|---|---|---|
| 11τ | 12τ | 13τ | 14τ | 15τ | 0τ |
| 15τ | 0τ | 1τ | 2τ | 3τ | 4τ |
| 3τ | 4τ | 5τ | 6τ | 7τ | 8τ |
| 7τ | 8τ | 9τ | 10τ | 11τ | 12τ |
| 11τ | 12τ | 13τ | 14τ | 15τ | 0τ |
| 15τ | 0τ | 1τ | 2τ | 3τ | 4τ |

**Figure 3.6:** Typical visual stimulation matrix of a c-VEP-based BCI. (A) Principle of equivalent neighbors. Sixteen target stimuli and twenty complementary stimuli are shown. (B) The time lag to the reference modulation code is indicated for each target.

Typically, the stimuli are displayed on standard computers. Similarly to the SSVEP paradigm, the stimuli (e.g., boxes containing letters) assume the binary states drawn/not drawn, which are updated every frame. The colors of the stimuli alternate between the foreground color (typically 'white', represented by '1') and the background color (typically 'black', represented by '0') in accordance with the used $m$-sequences.

The duration of one stimulation cycle is dependent on the monitor refresh rate, $r$. For example, if the monitor refresh rate is set to 60 Hz, the duration of one cycle of the flickering pattern is 63/60=1.05 s (see Figure 3.7), a time window that is reasonably short, but still long enough for reliable classifications. With a higher refresh rate the time for one stimulus cycle and the lag between consecutive targets decrease, which might impede the analysis of c-VEP responses.

In the BCI literature [9, 45, 46, 102, 103], $m$-sequences with a code length of 63 bit are the most popular choice for the stimulus presentation when employing the c-VEP paradigm. A code length of 63 bit is suitable for multi-target implementations. For example, if a time lag of 2 bit between adjacent stimuli is employed, a 63 bit $m$-sequence provides 32 stimuli, which is sufficient for spelling applications presenting the English alphabet, which contains only 26 letters and leaves room for six additional symbols (for example, German Umlauts [9]).

Wei et al. [103] investigated the impact of c-VEP stimulus specificity on the overall system performance. The parameters, size, color, proximity of the stimuli, and the length and lag of the stimulus sequence were tested. According to their results, a system with

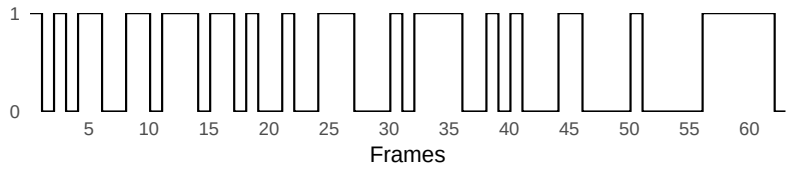



**Figure 3.7:** Stimulus pattern of a 63 bit $m$-sequence. Each bit corresponds to one frame. If the monitor refresh rate is 60 Hz (the refresh rate of a typical monitor), the stimulus duration of one cycle is 63/60 = 1.05 s.

white targets outperformed a system with red, green, blue, and yellow targets. Moreover, they found that the dimension of stimuli, the distance between stimuli, and the lag of the $m$-sequence should be selected as large as possible; furthermore, the code length should be as long as possible.

## 3.3 Classification - Minimum Energy Combination

Various factors, such as background processes in the brain, respiratory artifacts, movements of the electrical cable, and environmental noise, impede the interpretation of the brain signals via BCI. Friman et al. [104] proposed the MEC method for SSVEP signal processing, which determines a suitable weight set for the electrode signals with respect to noise cancellation.

### Model

To interpret the SSVEP response to a specific stimulus frequency $f$, we consider the following linear model, which describes the voltage $y_i(t)$ between the $i$-th electrode and a reference electrode at time point $t$ [42, 104]:

$$y_i(t) = \sum_{k=1}^{n_h} \big(a_{i,k} \sin\big(2\pi k f t\big) + b_{i,k} \cos\big(2\pi k f t\big)\big) + e_i(t). \quad (3.11)$$

The sum over the sine and cosine terms describes the SSVEP response to the stimulus frequency $f$, represented by $n_h$ harmonics; the set of coefficients $a_{i,k}$ and $b_{i,k}$ determines the amplitude and phase of the signal; the term $e_i(t)$ denotes the remaining information, i.e., the channel-specific noise and nuisance attributes to the signal.

This model can be generalized for multi-channel signals. We consider a data window of $n$ samples from $m$ electrode signals recorded at a sampling rate of $F_s$ Hz. The multi-channel EEG data are then represented as a matrix $\mathbf{Y} \in \mathbb{R}^{n \times m}$. For the sinusoidal terms, we construct a reference matrix $\mathbf{R} \in \mathbb{R}^{n \times 2n_h}$,

$$\mathbf{R} = \begin{bmatrix} \sin\big(2\pi f t\big) \\ \cos\big(2\pi f t\big) \\ \vdots \\ \sin\big(2\pi n_h f t\big) \\ \cos\big(2\pi n_h f t\big) \end{bmatrix}^T, \quad t = \frac{1}{F_s}, \frac{2}{F_s}, \ldots, \frac{n}{F_s}. \quad (3.12)$$

We can now generalize (3.11) to

$$\mathbf{Y} = \mathbf{RA} + \mathbf{E}, \tag{3.13}$$

where the coefficients corresponding to the sine and cosine entries of $\mathbf{R}$ are stored in the matrix $\mathbf{A} \in \mathbb{R}^{2n_h \times m}$, and the noise signals are stored in the matrix $\mathbf{E} \in \mathbb{R}^{n \times m}$.

### Design of Spatial Filters

The MEC uses principal component analysis (PCA) to find an optimal weight combination for the electrode channels. Such weight vector, $\mathbf{w} \in \mathbb{R}^m$, is then used to create a virtual channel, $\mathbf{s} \in \mathbb{R}^n$, holding the spatially filtered data,

$$\mathbf{s} = \mathbf{Yw}. \tag{3.14}$$

Several of such weights and virtual channels can be created: Let $n_s$ denote the total number of created virtual channels. By setting $\mathbf{S} = [\mathbf{s}_1, \ldots \mathbf{s}_{n_s}]$, we can generalize (3.14) to

$$\mathbf{S} = \mathbf{YW}, \tag{3.15}$$

where the columns of the weight matrix $\mathbf{W} \in \mathbb{R}^{m \times n_s}$ contain the weight vectors to the corresponding virtual channels, i.e., $\mathbf{W} = [\mathbf{w}_1, \ldots \mathbf{w}_{n_s}]$.

To determine the optimal weight matrix $\mathbf{W}$, the MEC approximates the noise matrix $\mathbf{E}$ from equation (3.13) and tries to minimize its energy. For this, an orthogonal projection is used to remove any SSVEP activity from the recorded signal; the signal matrix $\mathbf{Y}$ is projected onto the orthogonal complement of the reference model $\mathbf{R}$:

$$\mathbf{E} = \mathbf{Y} - \mathbf{R}(\mathbf{R}^T\mathbf{R})^{-1}\mathbf{R}^T\mathbf{Y}. \tag{3.16}$$

As the Matrix $\mathbf{E} \approx \mathbf{E}$ approximates the noise signal, the MEC searches a weight $\mathbf{w}$ that minimizes its energy,

$$\min_{\mathbf{w}} \|\mathbf{Ew}\|^2 = \min_{\mathbf{w}} \mathbf{w}^T\mathbf{E}^T\mathbf{Ew}. \tag{3.17}$$

To solve this optimization problem the eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_m$ and corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of the symmetric matrix $\mathbf{E}^T\mathbf{E}$ are determined.

As $\mathbf{E}^T\mathbf{E}$ is a symmetric matrix, the eigenvectors are pairwise orthogonal, and, in accordance with the Courant-Fischer min-max theorem (e.g., [105]), the quadratic form in (3.17) is bounded by the minimal and maximal eigenvalues. Hence, the solution of the

optimization problem is the eigenvector $\mathbf{v}_1$ corresponding to the smallest eigenvalue $\lambda_1$. The weight vectors can be selected as

$$\mathbf{w}_i = \frac{\mathbf{v}_i}{\sqrt{\lambda_i}}, \quad i = 1, \ldots, n_s, \tag{3.18}$$

yielding pairwise uncorrelated channels

$$\mathbf{s}_i = \mathbf{Y}\mathbf{w}_i, \quad i = 1, \ldots, n_s. \tag{3.19}$$

The number of virtual channels, $n_s$, can be selected arbitrary, as long as $1 \leq n_s \leq m$. As suggested by Friman et al. [104], the number of channels can be selected to filter out 90% of the noise signal. For this, $n_s$ is determined as the largest number such that

$$\frac{\sum_{i=1}^{n_s} \lambda_i}{\sum_{j=1}^{m} \lambda_j} < 0.1, \tag{3.20}$$

where the nominator represents the noise energy remaining in the $n_s$ virtual channel, and the denominator represents the total noise energy of the signal.

### Target Identification

The SSVEP signal energy $\hat{P}$ of a frequency $f$ and its $n_h$ harmonics in the spatially filtered channels $\mathbf{s}_1, \ldots, \mathbf{s}_{n_s}$ is calculated as

$$\hat{P} = \frac{1}{n_s n_h} \sum_{j=1}^{n_s} \sum_{k=1}^{n_h} \| \mathbf{R}_k^T \mathbf{s}_j \|^2, \tag{3.21}$$

where $\mathbf{R}_k \in \mathbb{R}^{m \times 2}$ refers to the sub-matrix of the frequency reference model $\mathbf{R}$ consisting only of the columns associated with the $k$-th harmonic,

$$\mathbf{R}_k = \begin{bmatrix} \sin\left(2\pi k f t\right) \\ \cos\left(2\pi k f t\right) \end{bmatrix}^T, \quad t = \frac{1}{F_s}, \frac{2}{F_s}, \ldots, \frac{n}{F_s}. \tag{3.22}$$

For signal classification, the MEC assumes all stimuli to be the target frequency and attempts to minimize the noise; virtual channels $\mathbf{s}_1, \ldots, \mathbf{s}_{n_s}$ are determined for all $K$ stimulation frequencies individually, as described in the previous section. To identify the actual target frequency, signal energies $\hat{P}_i$, $i = 1, \ldots, K$ corresponding to the individual frequencies are determined using (3.21) with the frequency-specific virtual channels. These energies are then normalized,

$$p_i = \frac{\hat{P}_i}{\sum_{j=1}^{K} \hat{P}_j}, \quad i = 1, \ldots, K. \tag{3.23}$$

The class label of the BCI-target $C$ is determined as the index of the frequency with the highest normalized signal energy,

$$C = \underset{i=1,\ldots,K}{\arg\max}\, p_i. \tag{3.24}$$

Asynchronous SSVEP applications can be realized by introducing threshold values for the classification. For example, the BCI output associated with the label $C$ may only be produced if $p_C$ surpasses a pre-set threshold, which is usually determined manually during a test session (see [42]).

To simplify the manual setup of these threshold parameters, Volosyak [42] suggested to increase the gap between the normalized signal energies of the stimulation frequencies by applying a softmax function:

$$p_i' = \frac{e^{\alpha p_i}}{\sum_{j=1}^{K} e^{\alpha p_j}}, \tag{3.25}$$

where $\alpha$ can be determined on the basis of the number of frequencies. By increasing $\alpha$, the gap between the normalized signal energies is enhanced. If $\alpha$ is too large, many misclassifications might occur. In many practical experiments where four or five frequencies were used, a value of $\alpha = 25$ resulted in high detection accuracies [42, 63]. The softmax function always outputs values between 0 and 1 with sum equal to 1. Instead of applying equation (3.24), the BCI-target $C$ is determined as the index maximizing $p_i'$.

## 3.4 Classification - Canonical-Correlation Analysis

CCA, introduced in 1935 by Hotelling [106], is an established method to investigate the relationships between two multi-dimensional variables. CCA finds a transformation of these variables, such that the transformed variables show maximal similarity to each other. The CCA can, therefore, be used to extract relevant information from the multi-channel EEG data while also reducing noise and nuisance signals. CCA is easy to implement and available in many programming languages, such as Python, R and MATLAB. Lin et al. [107] were the first to apply the CCA in the field of VEP-based systems. Meanwhile, CCA has also been applied to P300-based BCIs [108].

### Design of Spatial Filters

We consider two multi-dimensional variables $\mathbf{X} \in \mathbb{R}^{p \times s}$ and $\mathbf{Y} \in \mathbb{R}^{q \times s}$. CCA finds a pair of basis vectors (weight vectors) $\mathbf{w}_X \in \mathbb{R}^p$ and $\mathbf{w}_Y \in \mathbb{R}^q$, such that the correlation $\rho$ between the linear combinations $\mathbf{x} = \mathbf{X}^T\mathbf{w}_X$ and $\mathbf{y} = \mathbf{Y}^T\mathbf{w}_Y$ is maximized. The weights $\mathbf{w}_X$ and $\mathbf{w}_Y$ are found by solving

$$\max_{\mathbf{w}_X, \mathbf{w}_Y} \rho(x, y) = \frac{E[\mathbf{w}_X^T \mathbf{X}\mathbf{Y}^T \mathbf{w}_Y]}{\sqrt{E[\mathbf{w}_X^T \mathbf{X}\mathbf{X}^T \mathbf{w}_X] \cdot E[\mathbf{w}_Y^T \mathbf{Y}\mathbf{Y}^T \mathbf{w}_Y]}}, \tag{3.26}$$

where $E$ denotes the expectation operator. The value $\rho$ is the first, also called maximal canonical correlation. To construct further pairs $\mathbf{x}$, $\mathbf{y}$, the CCA searches additional weight vectors maximizing (3.26) subject to the restriction that they are uncorrelated with the first pair of canonical variables; this results in the second pair of canonical variables. This procedure can be repeated several times; in total, CCA can generate up to $\min\{p, q\}$ canonical correlations with corresponding weights.

For BCIs, the number of employed signal electrodes is usually more than twice the number of harmonics that are considered for signal classification ($m > 2n_h$). In this case, the maximal number of canonical correlations is $2n_h$.

Similarly to the MEC, which creates successive orthogonal eigenvectors, the first canonical variable yields the weight vector that yields maximum correlation. In most BCI-studies, and in this work, only the first canonical correlation is considered for the design of spatial filters and classification [107, 109]. With respect to the above, it should be noted that recent studies yielded significantly better results when employing additional correlations [110].

### Target Identification

We consider an SSVEP-BCI with $K$ stimulation frequencies $f_1, \dots f_K$. The EEG signals that the BCI interprets are stored in a signal matrix $\mathbf{Y} \in \mathbb{R}^{m \times n}$, where $n$ denotes the number of samples and $m$ the number of EEG channels.

For classification, we again use the sine and cosine reference signals defined in section 3.3, equation (3.12). For each of the $K$ reference signals $\mathbf{R}_1, \dots, \mathbf{R}_K$ constructed with stimulation frequencies $f_1, \dots, f_K$, the maximal canonical correlation with respect to the signal matrix $\mathbf{Y}$ is determined via CCA: Inserting $\mathbf{Y}$ and $\mathbf{R}_i^T$ in (3.26), yields $K$ canonical correlation coefficients $\rho_i$, $i = 1, \dots, K$. The

class $C$, associated with the BCI-target that the user is presumably focusing on is then determined, as

$$C = \underset{i=1,\ldots,K}{\arg\max}\, \rho_i. \tag{3.27}$$

## 3.5 Classification - Template Matching Method

For the c-VEP paradigm, a template matching approach can be used to identify the targets [9, 46]. For this approach, a training stage is required, where individual reference templates for all stimuli classes are created, and where spatial filters are designed. In a test stage, the spatial filters are applied to the reference templates and the test data (for example, data recorded in real-time on-line sessions). The spatially filtered templates are then individually compared to the spatially filtered test data.

The template matching method can also be used for the SSVEP paradigm. According to Nakanishi et al. [111], the waveforms of individual SSVEP EEG data and sinusoidal reference templates show consistent frequency components. As phase and amplitude of the fundamental and harmonic VEP responses vary among users, the collection of individual templates for the SSVEP paradigm could lead to higher classification accuracies. For the hybrid frequency and phase coding approach, the recording of individual EEG data is required to maintain the phase information. While some SSVEP classification methods incorporate both sinusoidal reference templates and individual EEG data (see, e.g., [66]), recent research yielded even better results, when only the individual EEG data were used for classification [52].

For the SSVEP paradigm, the spatial filters and templates need to be designed for each class individually. For the c-VEP paradigm, one spatial filter and one template are sufficient as the remaining targets can be generated via circularly shifting, as described in the following section. More details regarding the difference between the SSVEP and the c-VEP template generation are discussed in section 5.3 and section 5.4.

### Design of Spatial Filters and Generation of Templates

For c-VEP systems, the recording of several trials of only one target is sufficient as the stimuli patterns are circularly shifted versions of each other. In the recording stage, the user needs to focus on the reference target for several trials. A trial usually consists of the data collected during one stimulation cycle of the flickering stimulus. To determine the beginning of a trial, the data acquisition and the

stimulus presentation must be synchronized; section 3.6 provides more details regarding the synchronization. The recorded data are then segmented into single trials $\mathbf{T}_i \in \mathbb{R}^{m \times n}$, where $n$ denotes the number of samples per trial and $m$ the number of electrodes. Thereby we assume that the rows of these trials are centered, i.e., each channel has zero means. (This is achieved by subtracting the mean across channels of each channel individually.) These centered trials are then averaged resulting in a reference template $\mathbf{R}$,

$$\mathbf{R} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{T}_i, \tag{3.28}$$

where $N$ refers to the total number of recorded trials.

By shifting the columns of the generated template $\mathbf{R}$ circularly, templates $\mathbf{R}_k \in \mathbb{R}^{m \times n}$, $k = 1, \ldots, K$ associated with the remaining targets can be generated:

$$\mathbf{R}_k(i, j) = \mathbf{R}(i, j - (\tau_k - \tau_1)), \quad i = 1, \ldots, m;\ j = 1, \ldots, n, \tag{3.29}$$

where $\tau_k - \tau_1$ refers to the time lag in samples between the target corresponding to class $k$ and the reference target (here, target 1). Here, we assume that the actual time lags between targets are multiples of $1/F_s$; the bit lags between the code patterns and the sampling rate of the amplifier must be selected accordingly.

The training data can further be used to generate a CCA-based spatial filter $\mathbf{w} \in \mathbb{R}^m$, as suggested by Spüler et al. [54] and Bin et al. [46]. For this, two matrices are constructed with the training trials and the reference template,

$$\mathbf{T} = [\mathbf{T}_1 \mathbf{T}_2 \ldots \mathbf{T}_N] \quad \text{and} \quad \mathbf{R} = [\underbrace{\mathbf{R}\mathbf{R} \ldots \mathbf{R}}_{N}]. \tag{3.30}$$

These matrices are inserted into the CCA equation (3.26), which yields the weight vectors $\mathbf{w}_{\tilde{\mathbf{T}}}$ and $\mathbf{w}_{\tilde{\mathbf{R}}}$. The former, $\mathbf{w} = \mathbf{w}_{\tilde{\mathbf{T}}}$, is selected as a spatial filter, as it mimics the optimization of the correlation between single trials of unknown labels and reference templates.

### Target Identification

Pearson's correlation method (PCM) [112] can be used to classify a recorded EEG test data set of an unknown label. First, the data set is centered by subtracting the mean value in each channel individually. Let the resulting data be stored in a matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$, where $m$ represents the number of channels, and $n$ represents the number of samples. Applying the spatial filter $\mathbf{w}$, we get a one-dimensional vector $\mathbf{x} = \mathbf{X}^T \mathbf{w}$.

The Pearson sample correlation coefficient $\rho_k$ between the spatially filtered test dataset $\mathbf{x}$ and the spatially filtered template $\mathbf{r}_k = \mathbf{R}_k{}^T\mathbf{w}$ is determined as

$$\rho_k = \frac{\mathbf{x}^T\mathbf{r}_k}{\sqrt{\mathbf{x}^T\mathbf{x} \cdot \mathbf{r}_k^T\mathbf{r}_k}}. \tag{3.31}$$

The classified label $C$ is identified as the index that maximizes the correlation coefficient,

$$C = \underset{k=1,\ldots,K}{\arg\max}\, \rho_k. \tag{3.32}$$

## 3.6 Synchronizing Stimulus Presentation and Data Acquisition

For SSVEP-based BCIs using hybrid frequency and phase coding and for c-VEP-based BCIs, synchronization between the amplifier and stimulus presentation is required, because, in practice, the integer values used for the sampling frequency of the amplifier $F_s$ and for the monitor refresh rate $r$ are not precise, and small differences might accumulate over time. For synchronization, stimulus onset markers are typically sent to the EEG hardware. These time-stamps can be acquired using a photo-resistor or photo-diode attached to the screen [58, 113]. Another approach is to send the time-stamps from the stimulation computer to the amplifier using the parallel port [103].

In [72], we proposed a purely software-based approach, allowing the detection of stimulus onset without the need for additional hardware. Two timers were used to calculate the stimulus onset delay, $d_s$, which describes the time interval between the acquisition of EEG data blocks and stimulus onset.

The first time-stamp, $t_1$, was acquired directly after the flickering was initiated (in the thread dedicated to the stimulus presentation). The amplifier transmits the EEG data in blocks; the number of samples, $n_b$, of one block of EEG data is a variable of the amplifier software. The duration of the collection of one EEG data block in seconds is $d_b = n_b/F_s$. The second time-stamp, $t_2$, was acquired directly after receiving a block of EEG data (in the thread dedicated to signal classification).

The time interval between the start of data acquisition and stimulus onset can thus be calculated as $d_s = t_2 - d_b + t_1$. The number of samples collected before stimulus onset, $n_s$, can be determined as $n_s = [d_s F_s]$, where $[\cdot]$ denotes the rounding operator, which rounds to the nearest integer. Figure 3.8 illustrates the software-based synchronization approach.

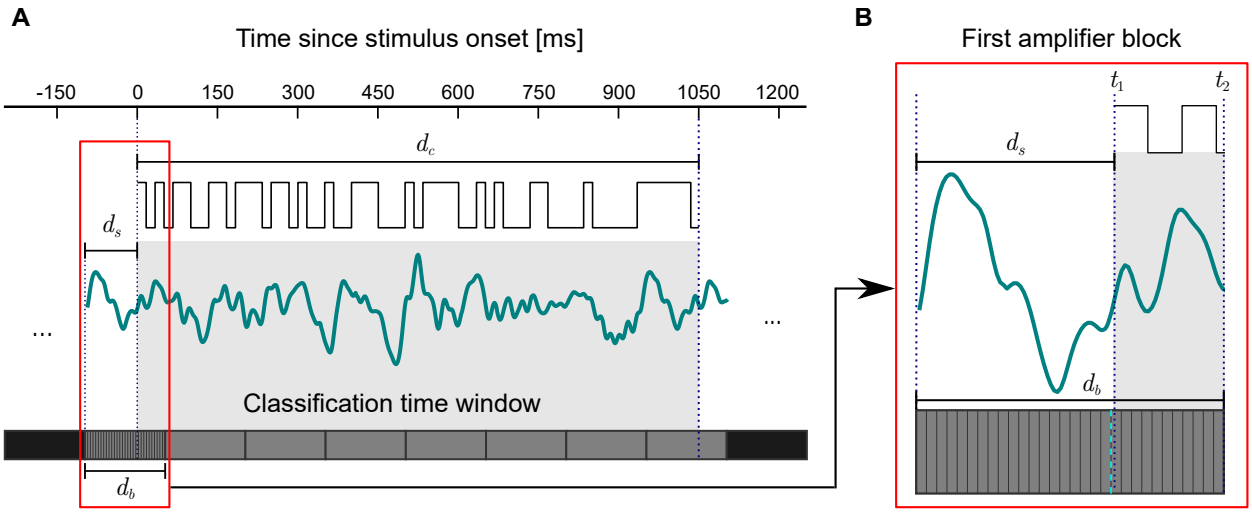



**Figure 3.8:** Software-based synchronization between signal acquisition and stimulus presentation. (**a**) The data collected during a full stimulus cycle (with a duration of $d_c$ ms) is shown. Displayed are the stimulus pattern and the EEG response recorded in one EEG channel. The amplifier sends the EEG data block-wise (i.e., in fixed time intervals of $d_b$ ms). The EEG-data collected prior to stimulus onset needs to be shuffled out. (**b**) The first amplifier block is shown. The time passed until stimulus onset, $d_s$, was determined after receiving the first block. It was calculated using the block duration, $d_b$, and two time-stamps ($t_2$ and $t_1$), which were set in the threads dedicated to the stimulus presentation and the signal acquisition, respectively. The dashed blue line indicates the last sample that is shuffled out. Figure adapted from Gembler and Volosyak [72].

## 3.7 Concluding Remarks

In general, BCI classification methods can be categorized into three groups with respect to the training aspect [114]:

**Training-free methods:**
These methods do not require any form of training data and can, therefore, be used immediately.

**Subject-independent training methods:**
For these methods, recorded data from various subjects are analyzed to determine fixed system parameters suitable for general users.

**Subject-specific training methods:**
These methods require the user-specific recording of EEG data; a training phase is used to optimize features and system parameters.

Examples of training free methods are the CCA and the MEC. Both of these methods can generally be used without any previous EEG recordings. The described template matching method, on the other hand, requires a recording session (subject-specific training).

Chapter 4 focuses on a training free application using the MEC (section 4.3). Some of the conducted experiments give insights to optimize fixed parameters such as the number of targets (section 4.5), addressing the second type of training (subject-independent optimization). A subject-specific calibration session for frequencies,

classification time windows, and thresholds, a so-called BCI wizard [115], will also be presented (section 4.4).

All c-VEP applications discussed in chapter 5 use subject-specific training in the form of personalized EEG data recordings.

# 4 Investigating SSVEP Parameters

Efforts to prevent BCI illiteracy when using the SSVEP paradigm usually include adjustments of key system parameters. This chapter provides insights into these critical parameters and their interactions.

The standard performance measure in BCI research, the ITR, depends on the average classification time, the classification accuracy, and the number of targets. Balancing these parameters ensures usability in regards to accuracy and efficiency. For maximal ITRs, the average classification time needs to be as low as possible, and the number of classes and the classification accuracy need to be as high as possible. In practice, however, these variables are not positively correlated. Long classification time windows lead to slow but accurate selections, and short time windows lead to fast but often undesired selections. For asynchronous systems, classification thresholds can balance speed and accuracy if carefully adjusted. Another essential factor in SSVEP-based systems is frequency selection. All of these parameters are dependent on the user; for example, some stimulation frequencies yield higher accuracies for some users. The question if optimal parameters could be anticipated with information about the user's age and gender is worth addressing.

In this chapter, several practical experiments conducted at Rhine-Waal University in order to investigate these parameters and their impact on BCI performance are summarized. In addition to that, methods to automatically determine these parameters using standard performance metrics are presented. It should be stressed that the optimization based on performance metrics does not address user-friendliness. The complexity of the GUI, initial orientation, and pleasantness of use are, however, critical points in regards to practical applications. In most of the reported experiments, questionnaires were conducted to address this aspect. The presented and discussed results were published before in journals and conference proceedings [61–63, 86, 116].

The chapter starts with a discussion of frequency selection (section 4.1) and classification time windows (section 4.2) for SSVEP systems. After that, the chapter investigates the impact of age on BCI performance. The multi-step spelling application used in the experiments summarized in this chapter is introduced, and the spelling performance of young and elderly users is compared (section 4.3). Following that, a wizard software that automatizes

the setup of critical parameters, namely frequencies, classification time windows, and thresholds, is presented. The summarized study about the wizard's functionality (61 participants) also investigates BCI performance differences between female and male subjects (section 4.4). The chapter ends with a report on two studies conducted to investigate the optimal number of stimuli for SSVEP-based BCI applications (section 4.5).

## 4.1 Impact of the SSVEP-Frequency Choice on SSVEP Performance

The period length of the stimulation frequency that the user is focusing on has a high impact on the strength of the evoked SSVEP response [43, 117]. The strength of the SSVEP response impacts the classification accuracy: The stronger the evoked SSVEP amplitudes, the more robust the target identification of the corresponding stimulation frequency.

In numerous studies, the relationship between frequencies and SSVEP amplitudes has been investigated to optimize BCI stimuli selection [6, 92, 118]. The SSVEP stimulus frequencies can be roughly divided into three ranges: Low frequencies up to 12 Hz (theta and alpha-band), medium frequencies between 12 and 30 Hz (beta-band), and high frequencies above 30 Hz (gamma-band) [26]. low and medium frequencies in the range of 6-20 Hz, i.e., the outermost part of the theta-band, the entire alpha-band, and the lower beta-band are the most popular choices for frequencies in SSVEP research [40, 52, 92]. For instance, the fastest SSVEP system tested to this date employed frequencies in the range from 8 to 15 Hz [52]. This range is suitable for multi-target systems, as overlapping effects between fundamental and harmonics are avoided.

Gao et al. [40] summarized the advantages of low-frequency stimuli for BCIs:

**Larger signal amplitudes:**
The amplitudes of SSVEPs elicited with low frequencies are higher in comparison to SSVEPs elicited with medium or high frequencies (see also [43]).

**No mutual influences:**
SSVEPs corresponding to medium and high stimulation frequencies may be misinterpreted as harmonic components of low-frequency stimuli or vice versa. If both low and medium or low and high frequencies are used simultaneously, the SSVEP response can not be uniquely matched to the

corresponding target stimulus. For example, a 16 Hz response could be elicited from both 8 Hz and 16 Hz stimuli.

**More robust against varying brightness:**
The SSVEP response that occurs when low stimulation frequencies are used is saturated more easily. The system is, therefore, robust if the surrounding brightness varies.

**Wider topographical distribution:**
The topographical distribution of brain responses evoked from low and high stimulation frequencies differ. For high stimulation frequencies, the topography of the SSVEPs is more restricted; for low stimulation frequencies, the SSVEPs distribute over a larger area of the head (see also [119]). Hence, if low stimulation frequencies are used, it is easier to find a suitable electrode configuration.

A disadvantage of the use of low stimulation frequencies is that users may perceive them as annoying and tiring [120, 121]. Visual fatigue of the user, especially when considering long-time use, limits the practicability of the BCI. Another disadvantage is a risk of photosensitive epileptic seizures triggered by the stimuli [121].

Because of these issues, many researchers tested high stimulation frequencies for BCIs [8, 118, 120, 122]. Although their SSVEP amplitudes are much weaker, some systems that employed high frequencies achieved promising results.

For example, Sakurada et al. [120] tested a BCI using three LEDs flickering at 61, 63, and 65 Hz. The participants achieved an average classification accuracy of 90% without experiencing visual fatigue. Chen et al. [118] also reached remarkably good results with high frequencies. In their study, the sinusoidal stimulation method was used to present 45 visual stimuli (ranging from 35.6 to 44.4 Hz) on a conventional LCD screen with a 120 Hz vertical refresh rate. A relatively high on-line average classification accuracy of 88.7% and an ITR of 61 bpm was achieved. Moreover, Volosyak et al. [8] compared medium-frequency range (13, 14, 15, and 16 Hz) to high-frequency range (34, 36, 38, and 40 Hz) using four LEDs as stimuli. In their study, 86 participants navigated a miniature robot through a labyrinth. The authors confirmed that BCI performance was influenced by the stimulation frequency. The medium-frequency setup could be controlled by 84 of 86 subjects with a mean ITR of 17.1 bpm and a mean accuracy of 92.3%; the high-frequency setup could be controlled by only 56 of 86 subjects with a mean ITR of 12.1 bpm and a mean accuracy of 89.2%. On the other hand, the authors stated that many subjects preferred the high-frequency BCI (although it yielded inferior speed and accuracy), as the flickering was less tiring for them.

## 4.2 Impact of the Classification Time Window on SSVEP Performance

In various studies, a strong correlation between BCI accuracy and the length of the time window dedicated to the SSVEP classification during EEG analysis has been observed [69, 123]. A short time window results in classification errors, but a long time window slows down the BCI performance.

In 2010, Volosyak et al. [123] addressed the relevance of the choice of appropriate time window length in a study with 10 participants: In a performance comparison on eight different time window lengths using five isolated frequencies (6.66, 7.5, 8.57, 10 and 12 Hz), the authors analyzed the distribution of the time window length for all correct classifications and reported an average time window length of 2.8 s for obtaining an SSVEP response recognition above 95%. The authors further observed that BCI control varied strongly across participants. This inter-subject variability was translated into notable speed differences between subjects. For instance, some users were able to successfully use the system with a time window as low as 0.5 s. Other users needed a time window of 3 s or more to achieve reliable control. The categorization of BCI performance with respect to demographic differences could lead to a better understanding of inter-subject variability. Some information about the user (e.g., age and gender) could be used to estimate optimal time windows and thus speed up the calibration.

Despite that, the BCI performance can also vary for one subject from one session to another (intra-subject variability or within-subject variance). Reasons for this kind of variability include lowered motivation, increased tiredness, or external distractions. Both intra-subject variability and inter-subject variability in classification time windows justify the use of asynchronous applications.

In asynchronous systems, the classification window length is closely linked to the classification threshold. If the classification threshold is not surpassed, further EEG data are collected; the classification window either increases (the new data are appended) or maintains its length (the new data are appended, and old data are shuffled out). The latter type of time window mechanism is referred to as sliding window (see, e.g., [42]). Sometimes, a combination of sliding and extending time windows is used [42, 63, 90]. The classification threshold needs to be calibrated carefully: If the threshold is set too low, misclassifications might be produced too often; on the other hand, if set too high, the system is slowed down too much. As the SSVEP amplitudes for different stimulation frequencies differ, the threshold should be set individually for each target. Indeed, in the aforementioned study, Volosyak et al. [123]

observed a difference in accuracy for different frequencies. For example, 12 Hz, the highest stimulation frequency in their study, yielded the weakest performance in comparison to the other stimuli for all tested time windows.

When performing a spelling task with a BCI speller, the user needs to shift his or her gaze between different stimulation frequencies. The transition phase from one stimulus to another can generate noise, which can interfere with the relevant EEG data (i.e., the data collected when gazing at the target). For this reason, gaze shifting phases, during which the flickering and data collection pause (see section 2.6), are integrated into these systems. Taking into account that the duration for frequency detection can be less than 0.5 s, the time for gaze shifting can be longer than the time for detection.

In this thesis, if not explicitly mentioned, the term classification window refers to the detection time only and does not include the gaze shifting phase, whereas the term selection time refers to the duration between two selections. For the calculation of the ITR, the latter is considered, i.e., the gaze shifting phase is always included. Some researchers employ gaze shifting phases as low as 0.5 s, which might be too low for untrained users (e.g., [52, 66]). In this thesis, gaze shifting phases of approximately 1 or 2 s are used.

## 4.3 Effect of Age on SSVEP Performance

This section is an amended version of [62]: Gembler et al. (2015), 'A Comparison of SSVEP-Based BCI-Performance Between Different Age Groups',
[63]: Volosyak et al. (2017), 'Age-Related Differences in SSVEP-Based BCI Performance'.

Severe disorders affect people from all age groups; the effects of aging alone present a range of physical limitations that prevent an interaction with the environment. Unfortunately, BCI prototypes are usually tested with young participants, typically students and employees of the research facilities. The age distribution is therefore strongly skewed to younger participants, roughly between 20 and 30 years. This age range is not representative of the general population; elderly participants are underrepresented.

A few field studies have been conducted to investigate demographic factors and their impact on BCI performance [8, 82]. In these studies, a trend that elderly users perform worse than young users was observed.

Several studies with smaller subject groups also suggest that BCI performance depends on the age of the users. Dias et al. [124] conducted a study with 12 participants to investigate the latency and distribution of P300. They found that elderly subjects (>51 years) show smaller P300 amplitudes than younger ones. Grosse-Wentrup and Schölkopf [125] reviewed performance variations in

SMR-BCIs and stated that a negative correlation between age and BCI performance is conceivable.

So far, very few studies investigated the effects of subject age with the SSVEP paradigm. Macpherson et al. [126] investigated age-associated changes in SSVEP amplitude and latency with memory performance. They found that older participants demonstrated reduced neural activity during lower task demands, whereas with greater task demands, their neural activity was increased. Ehlers et al. [127] reported age group distinctions concerning classification accuracies with an SSVEP-based spelling application; however, only children and young adults between 6 and 33 years were tested in this study. According to their findings, young adults obtained higher accuracy rates compared to children. Hsu et al. [128] studied the amplitude-frequency characteristics of frontal and occipital SSVEPs in young, elderly, and ALS patients. They found that the amplitudes of occipital SSVEPs in the young group (mean age 24.3 years) were significantly larger than the amplitudes of the elderly group (mean age 54.1 years). Norton et al. [129] reported that young children between 9 and 11 years could reliably control an SSVEP-BCI.

The mentioned field studies focusing on SSVEP-BCI demographics also reported age-associated performance differences. Allison et al. [82] analyzed the spelling performance of 106 participants (mean age 30.6 years, range 18-79) with the Bremen BCI spelling application. It was observed that younger subjects were less annoyed by the flickering and tended to attain higher ITRs; however, no statistical effect was found. In the subsequent demographics study [8], 86 subjects (mean age 25.8 years, range 18-55) were tested, but again, neither a statistically significant effect of age, gender, nor their interaction was observed.

Thus, the impact of age on the performance of SSVEP-based BCIs is worth further investigation. Based on the results of the studies mentioned before, we expected a negative correlation between age and BCI performance. The remainder of this section provides a summary of the experiment conducted in [63], where two equally sized groups with different age ranges (young and elderly) used an SSVEP-spelling application.

A critical factor in ensuring effective control is the arrangement and number of visual stimuli. Especially for older adults, the simplicity of the GUI and the readability of the letters are essential. In general, the BCI literacy rates and accuracies tend to be higher if a low number of targets is used (see section 2.6). For the presented experiment, an asynchronous four-target spelling application was designed. The impact of the number of targets on BCI performance will be addressed more thoroughly in section 4.5.

As discussed in the previous section, the choice of the classification time window is a crucial factor in ensuring reliable control. Therefore, in the presented study, selection time windows for the different age groups were investigated. Moreover, the standard measures (classification accuracy, ITR, and OCM) were analyzed.

## Methods

This section provides details about the subject group, the hardware, the software, and the experimental procedure. Furthermore, the design of the three-step spelling application, which uses four target stimuli, is presented. In terms of signal classification, the MEC was used (see section 3.3).

### Participants

Twenty subjects participated in the study. The participants were divided into two equally sized groups (i.e., ten subjects per group) according to their age.

Participants from the group of younger participants (in the following referred to as the young group) were recruited among the students of the Rhine-Waal University of Applied Sciences. Participants from the group of elderly participants (in the following referred to as the elderly group) were recruited among relatives from staff members and volunteers from nearby retirement homes.

Participants from the young group had a mean (SD) age of 22.4 (2.9) years, ranging from 19 to 27; four participants of this group were female. Participants from the elderly group had a mean (SD) age of 67.3 (5.7) years, ranging from 54 to 76; seven participants of this group were female.

All participants were BCI naïve (i.e., they had never used a BCI system before). All subjects had normal or corrected-to-normal vision. Spectacles were worn if needed.

Ethical principles were taken into consideration during all BCI experiments reported in this thesis. In this respect, all subjects gave written informed consent in accordance with the Declaration of Helsinki before participating. The research was carried out under best practice guidelines: The participants had the opportunity to withdraw from participation at any time; information needed for the analysis of the experiments was stored anonymously during the experiment; subjects questionnaire and signed consent forms were stored separately; results cannot be traced back to the participant.

The consent form included the following questions: "Do you understand that you are free to withdraw from the study at any time and without having to give a reason?", "Do you understand that your participation is completely voluntary, and if you do decide to be a subject, you may choose to leave at any time without penalty?"; the experiment was only conducted if the participant positively answered these questions and signed the consent form.

The entire session lasted approximately 60 minutes for each participant. The experiments were conducted in a typical laboratory setting with low background noise and luminance. The subjects did not receive any financial reward for their participation.

**Hardware**

Participants were seated at a distance of about 60 cm in front of an LCD screen (BenQ XL2420T, resolution: 1920 × 1080 pixels, vertical refresh rate: 120 Hz). The computer system operated on Microsoft Windows 7 Enterprise running on an Intel processor (Intel Core i7, 3.40 GHz).

Standard Ag/AgCl electrodes were used to acquire the signals from the surface of the scalp. Electrodes were mounted in accordance with the 10-5 system of EEG electrode placement (see section 2.2). The ground electrode was placed over AFz, the reference electrode over Cz (quite common locations for the ground and reference electrodes for BCI studies based on visual stimuli [18]). Eight signal electrodes were used; they were placed at the occipital region over the visual cortex: Pz, PO3, PO4, O1, O2, Oz, O9, and O10. Standard abrasive electrolytic electrode gel (Theodor-Körner-Apotheke, Graz, Austria) was applied between the electrodes and the scalp. The scalp was prepared by light abrasion until impedances were below 5 kΩ. An EEG amplifier, g.USBamp (Guger Technologies, Graz, Austria), was used with sampling frequency set to 128 Hz. During the EEG signal acquisition, a digital bandpass filter (between 2 and 60 Hz) and a notch filter (around 50 Hz) were applied.

**Sliding window mechanism for SSVEP signal classification**

An asynchronous threshold-based signal classification approach on the basis of the MEC was used (see section 3.3). Three additional pseudo-targets (selected as means between the four stimulation frequencies) were considered during classification. The recorded EEG-data were processed in blocks of 13 samples (approximately 0.1 s with the sampling rate of 128 Hz). An output command was only produced if the probability associated with the classified label met a threshold criterion. If no frequency probability exceeded

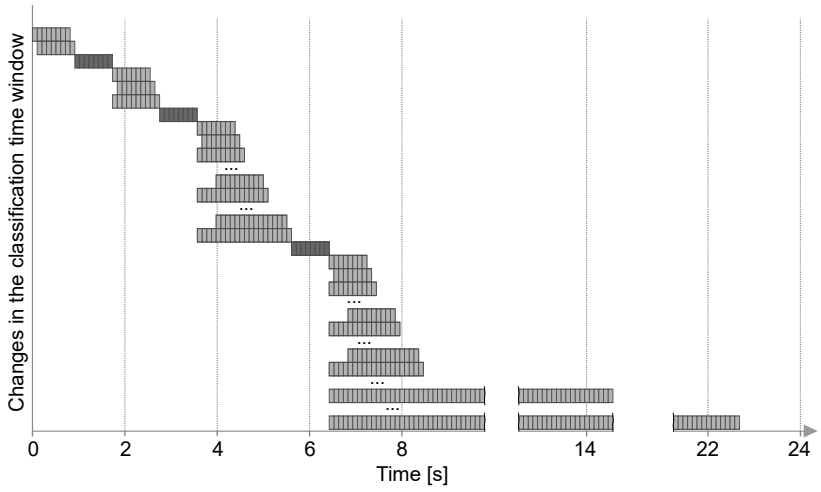



**Figure 4.1:** Dynamic classification times during on-line SSVEP spelling. If no output is generated and the actual time $t$ allows the extension to the next predefined value. After a performed classification, additional time for gaze shifting was included (dark gray). During this phase, the flickering paused, and no classification was performed. Figure modified from Gembler et al. [86].

the corresponding classification threshold $\beta_i$, or if one of the additional frequencies ($i > 4$) had the highest probability, the classification was rejected. For each stimulation frequency, the associated classification threshold was determined manually by the experimenters during the familiarization run (see experimental protocol, for more details).

When the threshold criterion was met, the system generated an output. Afterward, a gaze shifting period followed. During this period, the classifier output was rejected for the duration of 0.914 s (9 blocks), and the targets did not flicker.

Figure 4.1 illustrates the time window mechanism. The SSVEP signal classification was performed using sliding windows of stepwise increasing length $n_y$ (representing the number of recorded samples per channels) [42]. If the threshold criterion was not met, the time window slid, and only the newest data were used for classification. The EEG blocks were collected in several data buffers of different length that could store between 0.8125 s and 16.25 s of EEG data: $T_1 = 8 \cdot 13$, $T_2 = 10 \cdot 13$, $T_3 = 15 \cdot 13$, $T_4 = 20 \cdot 13$, $T_5 = 30 \cdot 13$, $T_6 = 40 \cdot 13$, $T_7 = 50 \cdot 13$, $T_8 = 60 \cdot 13$, $T_9 = 70 \cdot 13$, $T_{10} = 80 \cdot 13$ and $T_{11} = 160 \cdot 13$ samples. If the actual time $t$ allowed the extension of the classification window $n_y$ to the next predefined value $T_i$, this new value was used instead.

### Three-step spelling application

The three-step spelling application that was designed for this experiment allowed the selection of individual letters in three steps. It resembled earlier developed GUIs [42, 130]. The GUI presented four selection options to the user (see Figure 4.2).

At least three steps were necessary to choose a single letter. In the case of mistakes or misclassifications, the user needed to select the

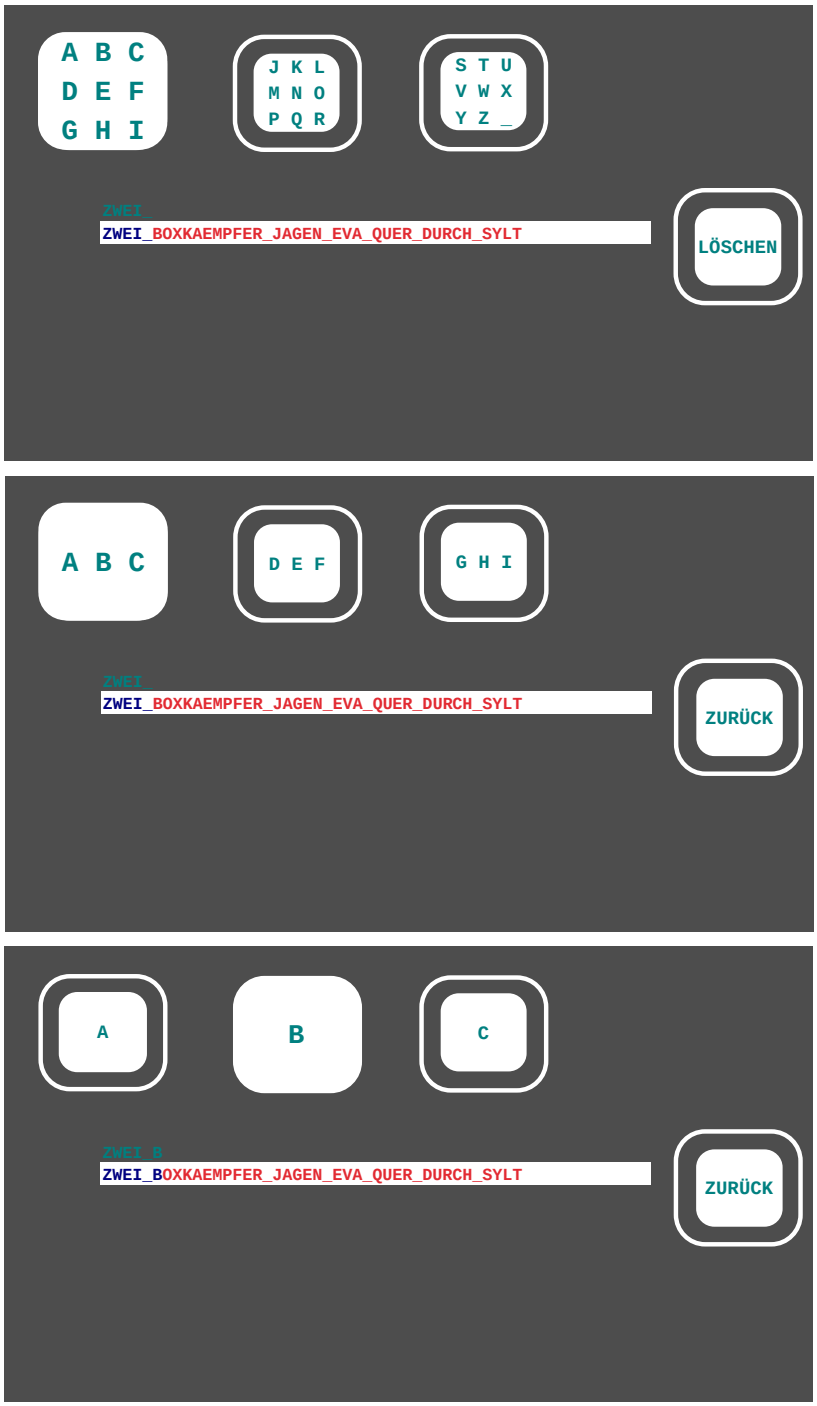



**Figure 4.2:** GUI of the three-step speller. A participant is writing the German pangram 'ZWEI_-BOXKAEMPFER_JAGEN_EVA_-QUER_DURCH_SYLT'. Individual letters could be selected in three steps. In the figure, the selection of the letter 'B' is presented.

correction option and the number of steps increased. In the first step, three boxes were arranged horizontally in the upper part of the screen containing the letters 'A-I', 'J-R', and 'S-_', respectively. A fourth box, positioned on the far-right of the screen, represented the correction option ('Löschen', the German word for delete). When the user selected this box, the last spelled character was removed from the written word.

In the second step, the content of the three boxes containing the alphabet changed to more specific sets. The boxes contained either

'A B C', 'D E F', 'G H I' or 'J K L', 'M N O', 'P Q R' or 'S T U', 'V W X', 'Y Z _', depending on the first selection. After selection in this second step, the content of the boxes changed to the individual letters, e.g., A, B, and C, if 'A B C' was selected in the second step. In both the second and the third step, the far-right box contained the command 'Zurück' (back), which allowed the user to return to the previous step. To reduce the cognitive load in the visual channel, the BCI provided audio feedback voicing the selected command or letter (also in German). The size of the boxes varied in relation to the SSVEP power distribution during the spelling (as described in [42]), providing continuous feedback of the classifier state. The default size of the boxes was 175 × 175 pixels. The boxes were outlined by a frame which determined their maximum size. The GUI presented the texts of the output word and the spelling task in the center of the screen.

#### Experimental protocol

After signing the consent form, each participant went through a pre-questionnaire, answering questions regarding gender, age, and previous BCI experience. Afterward, the subjects were prepared for the EEG recording. The stimulation frequencies were generated with dividers of the refresh rate (120 Hz), as described in section 3.1. Subjects participated in a familiarization run where the words 'BCI', 'KLEVE', and a word of choice (e.g., the own first name) were spelled. If repeated false classifications occurred during this test run, the experimenters manually adjusted the classification thresholds or chose different frequencies (all between 6 and 12 Hz). After the familiarization run, each subject used the GUI to spell the German pangram 'ZWEI BOXKAEMPFER JAGEN EVA QUER DURCH SYLT'. The flickering stopped automatically when the participant finished the copy spelling task. Spelling errors needed to be corrected via the implemented delete function. After the on-line spelling task, the subjects completed a brief post-questionnaire.

### Results

Differences in BCI performance between the young and elderly participants were assessed by comparing the classification accuracy, ITR, and OCM achieved with the pangram task. Figure 4.3 A shows the individual accuracies and ITRs for the young and elderly group, respectively; Figure 4.3 B compares the classification accuracy, ITR, and OCM of the two tested age groups.

For both groups, all participants were able to complete the spelling task achieving reliable control, reaching accuracies above 70%;

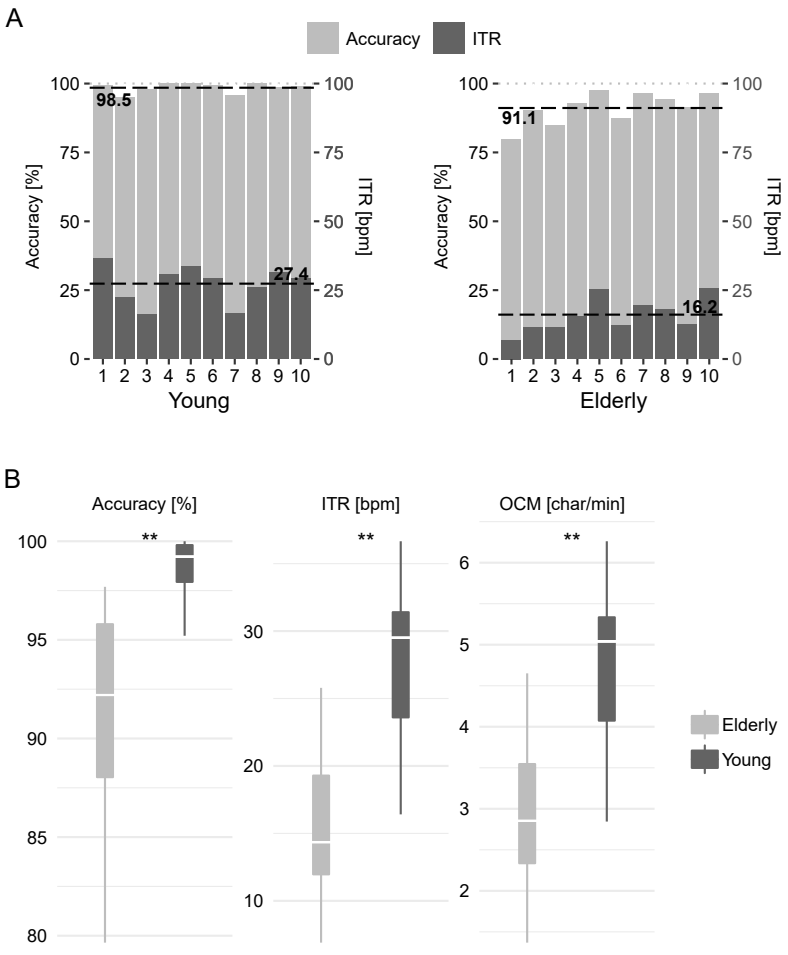



**Figure 4.3:** Comparison of SSVEP on-line spelling performances of elderly and young participants. The task was to spell the sentence 'ZWEI BOXKAEMPFER JAGEN EVA QUER DURCH SYLT' (a German pangram). (A) Individual ITRs and accuracies of the ten participants per age group. The dashed lines indicate the mean values. (B) Compared are ITR, classification accuracy, and OCM for the two different age groups. The significance of Welch's $t$-test are marked by asterisks; * indicates $p < 0.05$, and ** indicates $p < 0.01$.

most participants achieved accuracies above 90%. The accuracies varied among participants (range 95-100% and range 80-98% for the young and elderly participants, respectively). Subjects from the young age group reached a mean (SD) accuracy of 98.5 (1.7)%. Three subjects from this group completed the spelling task without errors, achieving an accuracy of 100%. Subjects from the elderly group reached a mean accuracy of 91.1 (5.4)%, and no subject of this group reached 100% accuracy. A Welch's $t$-test (also referred to as unequal variances $t$-test) revealed a significant difference between the mean accuracies of young and elderly subjects ($t = 3.88$, $p = 0.002$).

In respect to the OCM measure, the young age group achieved, on average, 4.7 char/min (SD 1.1, range 2.8 - 6.3 char/min). The elderly group achieved, on average, 3.0 char/min (SD 1.1, range 1.4 - 4.6 char/min). For this measure, the difference between the groups was also significant ($t = 3.36$, $p = 0.003$).

The number of targets of the three-step speller, $N = 4$, was used to determine the ITR (calculated as described in 2.7). The lower bound of the average time of a selection is the sum of the duration of the gaze shifting phase (0.914 s) and the minimal classification time window of the detection algorithm (0.813 s). Inserting $N = 4$,

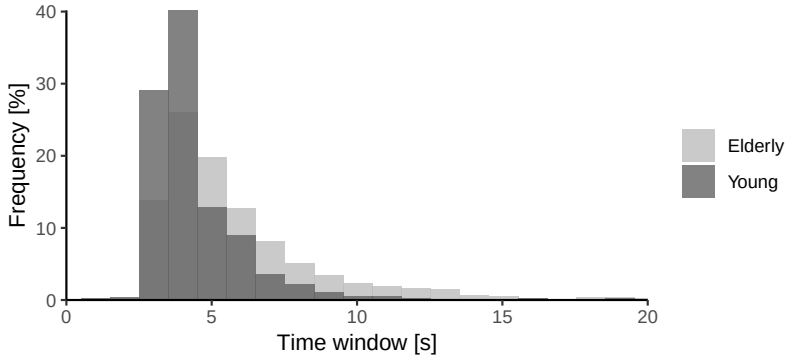



**Figure 4.4:** Selection time windows for elderly and young participants. The distribution of selection time windows for the on-line-experiment 'ZWEI BOXKAEMPFER JAGEN EVA QUER DURCH SYLT' is displayed.

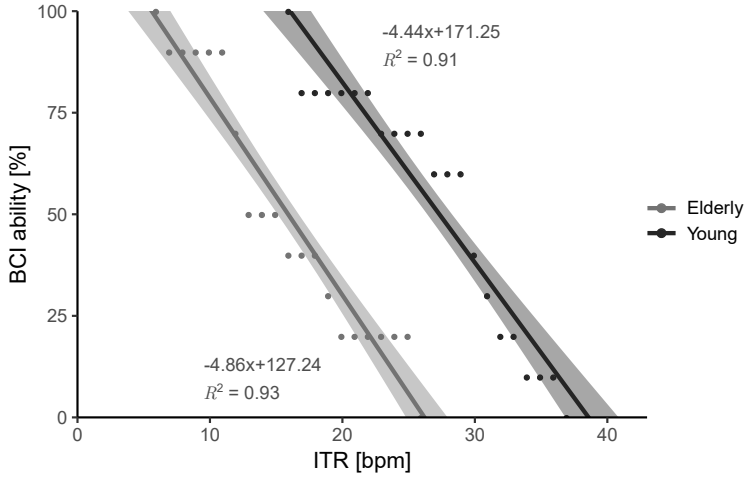



**Figure 4.5:** SSVEP-BCI ability for the three-step speller in relation to the achieved ITRs for young and elderly participants. The gray band indicates the 95% confidence interval around linear regression lines for the two groups.

$t = 1.727$, and $p = 1$ into the ITR formula yields a theoretical maximal ITR of 69.5 bpm.

In practice, however, the average times for correct commands were much longer than the theoretical minimum value: Young participants needed 4.4 s and elderly participants needed 6.5 s on average. The difference between the groups was significant ($t = 2.98$, $p = 0.008$). Figure 4.4 shows the distribution of classification times for all selections in the experiment for young and elderly users. It can be seen that for elderly users, the BCI more frequently used long classification time windows (>5 s) to produce outputs.

The achieved ITRs were lower than the theoretical ITR as well. Subjects from the young group achieved a mean ITR of 27.4 (6.5) bpm; subjects from the elderly group achieved a mean ITR 16.2 (5.9) bpm. The difference in ITR between the groups was again significant ($t = 3.85$, $p = 0.001$).

Furthermore, Figure 4.5 displays the BCI ability (see section 2.8). Each dot in the figure presents the percentage of young and elderly participants, respectively, who achieved or surpassed the corresponding ITR on the $x$-axis in the on-line copy spelling task. The data from the 10 subjects for each age group were fit into linear models to generalize the percentage of young or elderly users

**Table 4.1:** Questionnaire results. The numbers are represented as number of respondents. The subjective impressions regarding fatigue level were measured using a five-point Likert scale [131]: (1) not tired, (2) little tired, (3) moderately tired, (4) tired, (5) very tired. Data from Volosyak et al. [63].

| | Level of tiredness | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Young | | | | | Elderly | | | | |
| | (1) | (2) | (3) | (4) | (5) | (1) | (2) | (3) | (4) | (5) |
| Before the experiment | 2 | 2 | 6 | 0 | 0 | 3 | 4 | 2 | 1 | 0 |
| After the experiment | 2 | 2 | 6 | 0 | 0 | 2 | 1 | 6 | 1 | 0 |

| | Flickering was annoying | | | |
|---|---|---|---|---|
| | Young | | Elderly | |
| | Yes | No | Yes | No |
| After the experiment | 5 | 4 | 4 | 6 |

that can achieve specific ITR values for the tested application. For example, based on these models, it is expected that about 82% of young and about 30% of elderly users can reach an ITR of at least 20 bpm. As noted, the theoretical maximal ITR with this system is 69.5 bpm. However, following the model for young subjects, the max ITR is around 40 bpm; i.e., in practice, the theoretical value is not achievable for untrained users.

Table 4.1 summarizes the questionnaire replies regarding tiredness and level of annoyance from the pre- and post-questionnaire.

### Discussion

Significant differences between the BCI performance of young and elderly users were observed. Results based on twenty healthy subjects demonstrated that thanks to the implementation of large classification time windows (up to 16 s), every subject gained control over the system. Commands were classified faster and more accurately for subjects of the young group. The results reveal that subject age influences BCI performance significantly and indicate that the subject age needs to be considered when designing GUIs and calibrating system parameters. The knowledge of the user age certainly helps to determine critical parameters for the BCI.

In the presented study, the classification time windows for subjects from the elderly group were usually larger (see Figure 4.4). Generally, the study confirms that the implementation of larger time windows is beneficial for some users.

A reason for the performance difference between the two age groups could be smaller SSVEP amplitudes of older adults, as

observed in many other studies, e.g., in the study of Hsu et al. [128]: The authors measured SSVEPs induced by different stimulation frequencies (ranging from 13 to 31 Hz) measured from both the occipital region and the frontal region. Eight young participants with a mean (SD) age of 24.3 (2.43) years and eight elderly participants with mean (SD) age of 54.1 (1.96) years were tested. They found that for all stimulation frequencies, the elderly group reached lower mean SSVEP amplitudes than the young group.

Another explanation for the performance difference could be shorter reaction times of the younger participants with respect to the spelling interface.

The performance gap might increase even further if a higher number of stimulation targets is used. Older people might have more problems with the increased information load on the visual channel. Low target systems offer more freedom in stimulus size, allow for a greater distance between stimuli, and require less precise gaze direction. As discussed in section 4.1, an important issue regarding user comfort in SSVEP-based BCIs is frequency selection. When asked about the discomfort caused by the flickering, 45% of the subjects stated that they found the flickering annoying; four of the elderly subjects even reported a slightly increased level of tiredness after the experiment (see Table 4.1).

It is well known that high frequencies produce less visual fatigue than lower frequencies and show no stimulus-related seizures (see section 4.1). However, the age-related performance gap might increase when using higher flickering rates. This issue will be investigated in chapter 5 with the c-VEP paradigm.

## 4.4 Automated Calibration for SSVEP-based BCIs

This section is an amended version of [86]: Gembler et al. (2015), 'Autonomous Parameter Adjustment for SSVEP-Based BCIs with a Novel BCI Wizard'.

Automated calibration phases to obtain subject-specific system BCI parameters could yield higher literacy rates. In the study summarized in this section, we presented a wizard, that determines critical SSVEP parameters for each subject individually. Many researchers emphasized the importance of practical BCI prototypes that will have an impact on the life quality of disabled people (e.g., [132]). Usability challenges have impeded the use of BCIs in everyday scenarios for a long time. Extensive effort has been made to conduct practical studies with the intended end-users [133–137]. Conducting and validating studies with the target population is much more challenging than proof-of-concept prototype tests in research facilities. Due to the preparation and setup, BCI experiments are

time intensive. Some of the studies with end-users highlight the importance of expert-independent system setup.

For example, Sellers et al. [137] tested a BCI used by a 51-year-old ALS-patient at his home. The system was used successfully for over 2.5 years and restored the user's independence in social interactions. Recalibration was performed remotely (over the Internet). Holz et al. [133] also installed a BCI controlled application at a locked-in ALS-patient's home. The authors demonstrated expert-independent home-use in their study, but they also reported varying performance and stressed the importance of regular calibration. In terms of practical solutions for clinical applications using SSVEP-BCIs, Punsawad and Wongsawat [138] proposed a system that can be enabled or disabled by alpha-band EEG and thus requires less assistance from the caregiver.

SSVEP systems rely on a variety of different parameters that influence BCI performance. As discussed in the previous sections, optimal stimulation frequencies and classification time windows can vary between users and need to be adjusted precisely to achieve reliable performance.

Such precise parameter set up cannot be expected from users or caregivers. Therefore automatized calibration methods are an essential step for BCIs to progress from laboratory demonstrators to practical real-life applications.

It should be noted that calibration methods have already become standard for P300-based BCIs. Typically, in a supervised classifier, EEG data are recorded in a calibration phase during which the user is asked to perform specific tasks. The recorded brain signals are then analyzed and decoded to customize control parameters. For example, Kaufmann et al. [75] developed a user-centered P300-BCI application that adjusts classifier weights and control parameters individually in the background. At the same time, research also focuses on so-called zero-training BCIs with shortened or omitted calibration periods [35, 139, 140]. Although a wizard that sets up essential parameters for the SSVEP-paradigm was suggested already in 2010 [92], calibration software for SSVEP-based BCIs has rarely been reported. An explanation for this is that many research groups use synchronous systems that require less setup (e.g., no calibration of thresholds is required) but might be less accurate and intuitive. In the study presented in this section, a wizard, that allows parameter setup with one click, was designed and tested with 61 participants. We further explored BCI demographics based on the data of this comparably large number of participants. BCI studies of this size have been conducted with the P300 [37], the MI [141], and the SSVEP paradigm [8, 82, 142]. Most of these studies

reported subjects that were not able to gain satisfactory control over the system, i.e., BCI illiteracy.

In spite of the ongoing incremental improvements in software algorithms, the BCI illiteracy phenomenon remains a reoccurring problem in SSVEP studies [1, 8, 82].

The overall aims of this research were:

- to investigate optimal stimuli selection for SSVEP-based BCIs through analysis of the wizard outputs;
- to show that the vast majority, if not all BCI users can control an SSVEP-based BCI application; and
- to prove that generally higher classification accuracies can be achieved (through autonomous parameter adaption by the wizard).

To demonstrate the functionality of the presented wizard software, we conducted an on-line copy spelling experiment after autonomous parameter setup.

## Methods

This section summarizes the methods and materials of the experiment. The focus is put on the presented wizard software. The hardware setup and signal processing methods were identical to the study presented in the previous section 4.3. The MEC was used for signal classification (see section 3.3), and the stimulus presentation was based on dividers of the vertical refresh rate (see section 3.1). For the copy spelling task, we used the three-step spelling interface with minor modifications.

### Participants

All 61 subjects (healthy adult volunteers) gave written informed consent following the Declaration of Helsinki. The participants had a mean (SD) age of 22.8 (5.0) years (range 17-49). The impact of gender on BCI performance was also evaluated in this study; 17 participants were female, 44 participants were male. Female participants had a mean (SD) age of 22.7 (4.4) years (range 18-36). Male participants had a mean (SD) age of 22.8 (5.3) years (range 17-49).

All subjects were students or employees of the Rhine-Waal University of Applied Sciences. The EEG recording took place in a standard laboratory room with low background noise and luminance. None of the subjects had neurological or visual disorders. Spectacles were worn when appropriate. Subjects did not receive any financial reward for participating in this study.

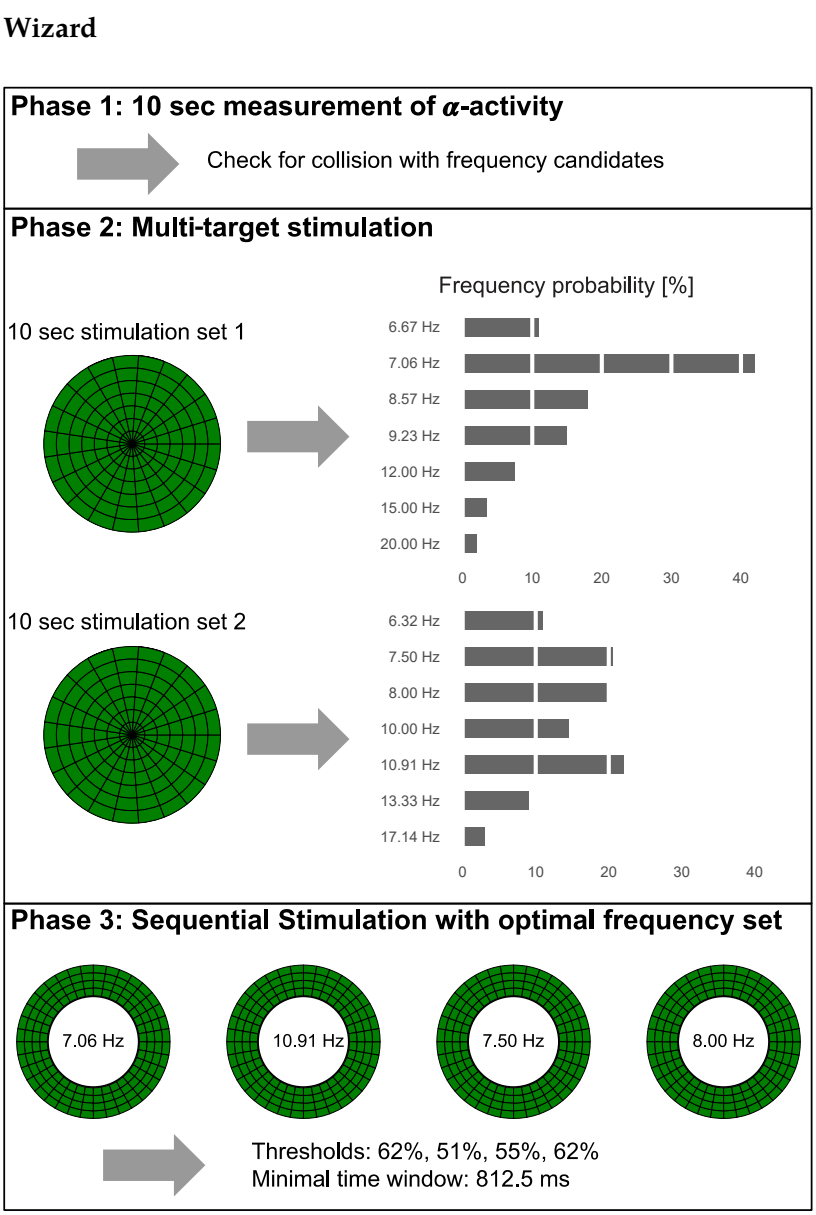



**Figure 4.6:** Illustration of the wizard's calibration procedure. In phase 1 (alpha test), EEG data were recorded (with the user's eyes closed) and analyzed to determine the alpha frequency. In case the alpha frequency overlapped with a stimulation candidate, this stimulus was neglected in the further procedure. In phase 2 (multi-target stimulation), EEG-data were recorded while the subject faced two circles containing stimulation candidates in sequence (each flickered for 10 s). The four frequencies yielding the strongest signal power were used as SSVEP targets. In phase 3 (sequential stimulation), participants gazed at each of the four determined frequencies individually. With the recorded data, SSVEP key parameters were determined.

The wizard ran the user through three phases in order to provide subject-specific stimulation frequencies (phases 1 and 2), classification thresholds, and classification time windows (phase 3). Figure 4.6 illustrates the entire calibration procedure for one subject. The techniques used in each step were derived from several previous findings. The so-called multi-target technique for the selection of individual subject-dependent stimulation frequencies, presented by Volosyak et al. [92], was based on the dual stimulation technique suggested by Mukesh et al. [143] where frequency combinations were used to increase the number of SSVEP targets.

The first task of the wizard was to select four optimal stimulus frequencies. Because of the advantages of low-frequency stimuli for SSVEP-based BCI (see section 4.1), the wizard was restricted to frequencies up to 20 Hz. As the number of suitable frequencies on the LCD monitor was limited by the vertical refresh rate of 120 Hz,

the wizard took only fourteen candidate stimuli into consideration. More specifically, the four optimal frequencies were determined from the frequencies 6.32, 6.67, 7.06, 7.50, 8.00, 8.57, 9.23, 10.00, 10.91, 12.00, 13.33, 15.00, 17.14, and 20.00 Hz (obtained with dividers of the refresh rate between 6 and 19, see Table 4.2).

**Classification method** As in the experiment summarized in section 4.3, to increase system robustness, three additional pseudo-targets, selected as means between two target frequencies, were considered for classification.

The MEC classifier output, $O$, was set to the index of the $i$-th frequency if the following conditions were met:

- the $i$-th frequency had the highest probability $p_i$;
- the detected frequency was one of the stimulation frequencies (i.e., none of the pseudo-targets had the highest probability); and
- the probability $p_i$ exceeded a predefined threshold $\beta_i$.

If no frequency probability exceeded the corresponding threshold $\beta_i$ or if one of the additional frequencies had the highest probability, the output label $O$ was set to zero.

EEG-data were processed by the computer in blocks of 13 samples (sampling rate 128 Hz). For classification, these EEG blocks were collected in several data buffers of different length $T_i$ (for more details, see section 4.3). The following pre-set time windows were used: $T_1 = 8 \cdot 13$, $T_2 = 10 \cdot 13$, $T_3 = 15 \cdot 13$, $T_4 = 20 \cdot 13$, $T_5 = 30 \cdot 13$, $T_6 = 40 \cdot 13$, $T_7 = 50 \cdot 13$, $T_8 = 60 \cdot 13$, $T_9 = 70 \cdot 13$ und $T_{10} = 80 \cdot 13$ samples.

| Period length | 19 | 18 | 17 | 16 | 15 | 14 | 13 |
|---|---|---|---|---|---|---|---|
| Frequency [Hz] | 6.32 | 6.67 | 7.06 | 7.50 | 8.00 | **8.57** | **9.23** |
| Period length | 12 | 11 | 10 | 9 | 8 | 7 | 6 |
| Frequency [Hz] | **10.00** | **10.91** | **12.00** | 13.33 | 15.00 | 17.14 | 20.00 |

**Table 4.2:** List of suitable SSVEP frequencies that can be generated with a refresh rate of $r = 120$ Hz. The frequency is calculated as $r/L$, where $L$ denotes the number of frames in a period. The frequencies overlapping with the alpha-band are marked bold.

**Phase 1 (Alpha test).** In the first phase of the wizard, it was tested if high alpha wave activity occurred. The low-frequency band overlaps with the alpha-band, which may be a cause for false classifications (see section 4.1). The candidate frequencies (see Table 4.2) were tested for interference with the user's alpha wave. This was realized as follows. After the wizard program was started, the user was instructed to close his or her eyes (by an audio instruction and a text message displayed on the screen). During the closed eye period, EEG data were recorded. After 10 s, a second audio message instructed the user to open his or her eyes again.

Alpha frequency interference was tested as follows. Using the MEC, averaged probabilities for the five stimulation frequencies $f_{st}$=8.57, 9.23, 10.00, 10.91, 12.00 Hz (all frequency candidates that belong to the alpha-band) and ten neighboring frequencies $f_{st}\pm 0.3$ Hz were generated on the basis of the recorded data.

For these frequencies, the MEC probability distributions for different classification time windows were calculated and averaged. More specifically, the power distributions for $p_{i,T_j}$ were calculated for all classification time windows and for all frequencies (e.g., after receiving 8 blocks, a value for $p_{i,T_1}$ was determined). In total, 100 blocks of EEG-data were recorded. For all time windows of lengths $T_j$, the MEC probabilities were averaged. For example, 100-8+1=93 overlapping blocks, each containing $T_1$ samples per channel, were analyzed and an averaged probability $\bar{p}_{i,T_1}$ was determined. Finally, the mean value over all averaged probabilities was calculated,

$$\bar{p}_i = \frac{1}{10}\left(\bar{p}_{i,T_1} + \bar{p}_{i,T_2} + \ldots + \bar{p}_{i,T_{10}}\right). \tag{4.1}$$

If one of the possible target frequencies had the highest averaged probability and surpassed a threshold of 0.1, it would be further on neglected. As a result, all of the remaining frequencies differed from the alpha wave by 0.15 Hz or more.

**Phase 2 (Multi-target stimulation).** In the second phase, multi-target stimulation was used to determine a set of four frequencies out of the fourteen suitable stimulation frequencies (see Table 4.2). For this frequency selection, the user faced a circle that represented seven of the candidate frequencies. The circle (radius 245 pixels) was divided into 147 segments (seven rings, each containing 21 segments). Each of these segments flickered during the recording, and each of the seven candidate frequencies was represented by 21 segments, which were scattered randomly.

After subjects were instructed by an audio message to shift their gaze on the circle, the flickering started, and EEG data were collected for approximately 10 s (100 blocks of EEG-data). Thereafter, the flickering paused for 2 s. The user then faced a second circle, which represented the seven remaining candidate frequencies. EEG data were recorded for another 10 s.

Each circle represented a mix of higher and lower frequencies. The frequencies represented in the first circle were 6.32, 7.50, 8.00, 10.00, 10.91, 13.33, and 17.14 Hz; the frequencies represented in the second circle were 6.67, 7.06, 8.57, 9.23, 12.00, 15.00, and 12.00 Hz. For each circle, to avoid mutual influences between frequencies, the additional restrictions rules $f_i \neq [f_j + f_k]/2$, $f_i \neq 2f_j - f_k$ (see [8]) were satisfied.

| $\beta_j$ [%] | $p_{\text{correct}}$ [%] | $p_{\text{false}}$ [%] | $p_{\text{zero}}$ [%] |
|---|---|---|---|
| 30 | 97.5 | 0 | 2.5 |
| 31 | 96.3 | 0 | 3.7 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| 50 | 44.4 | 0 | 55.6 |
| **51** | **43.2** | **0** | **56.8** |
| 52 | 39.5 | 0 | 60.5 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| 69 | 23.5 | 0 | 76.5 |
| 70 | 21.0 | 0 | 79.0 |

**Table 4.3:** Example of the threshold determination of the wizard software. Provided are the distributions of classifier outputs $p_{\text{correct}}$, $p_{\text{false}}$, and $p_{\text{zero}}$ for a fixed frequency $i$ and a fixed time window of length $T_s$. In this example, the threshold for the corresponding frequency was set to 51%.

If one of the fourteen frequency candidates interfered with the user's alpha wave, this frequency was left out. In that case, one of the circles represented only six frequencies (it had only six rings).

After the data for both circles were recorded, for each circle, the probabilities (4.1) of the represented frequencies were determined. Thereafter, the fourteen candidate frequencies were ranked from the highest averaged probability to the lowest. The four highest-ranked frequencies were selected as suitable target frequency. However, the four selected frequencies needed to meet the restriction rules. If they were not met, the lowest-ranked of the frequencies causing the violation was replaced by the highest-ranked frequency from the remaining stimuli. For example, if 6 Hz had the highest and 12 Hz the second-highest averaged probability, the latter was replaced, as otherwise the restriction rules were violated.

**Phase 3 (Sequential stimulation).** In the third phase of the wizard, optimal classification thresholds were determined for each of the four frequencies determined in the previous step. The user needed to gaze at these frequencies in sequence for 10 s each. The target frequency was represented by a white circle (radius 150 pixels), initially flickering at the frequency, which was ranked highest in phase 2.

When gazing at a target in a typical SSVEP-BCI application, the neighboring stimuli in the user's peripheral vision add to the noise signal. To simulate this noise, four green rings (144 segments in total, outer diameter 500 pixels) surrounded the white circle (representing the target frequency). Within these rings, each of the remaining three frequencies from the determined frequency set was represented by 144/3=48 segments, which were scattered randomly. Before the flickering and the recording started, the user was instructed by an audio message to gaze at the inner white circle.

The circle and the surrounding rings flickered for 10 s while EEG data were recorded. Then the flickering paused for 2 s. This pause ensured that the SSVEP-responses from the first recording did not influence the following recording. After that, the white circle represented the second-ranked frequency from phase two, and the rings represented the remaining three frequencies accordingly. This procedure was repeated until individual EEG recordings for all four optimal frequencies were collected. In total, the recording time in phase 3 was 40 s.

To determine the optimal classification thresholds and the optimal time window the MEC classifier outputs $O_i(T, \beta)$ of the individual frequency recordings ($i = 1, 2, 3, 4$), were determined for all preset time windows $T = T_1, \ldots, T_{10}$ and the suitable thresholds $\beta = 0.30, 0.31, \ldots, 0.69, 0.70$. The outputs $O_i(T, \beta)$ were then categorized into three classes:

- if $O_i(T, \beta) = i$, i.e., if the output was equal to the index of the stimulation frequency, it was categorized as 'correct classification';
- if $O_i(T, \beta) \neq i$ and if $O_i(T, \beta) > 0$, i.e., if the output was equal to the index of one of the remaining three stimulation frequencies, it was categorized as 'false classification'; and
- if $O_i(T, \beta) = 0$, i.e., if no frequency probability exceeded the threshold $\beta$ or if an additional frequency had the highest probability, the output was categorized as 'zero classification'.

For each stimulation frequency ($i = 1, 2, 3, 4$) the distributions of correct classifications, $p_{\text{correct}}(i, \beta, T)$, false classifications, $p_{\text{false}}(i, \beta, T)$, and zero classifications, $p_{\text{zero}}(i, \beta, T)$, were determined.

The frequency-specific optimal thresholds $\beta_i$ and the optimal minimal time window lengths were determined in an iterative process: First, for each frequency, the largest threshold value $\beta = 0.30, 0.31, \ldots, 0.70$, which satisfied the conditions

$$p_{\text{correct}}(i, \beta, T) \geq 0.4 \text{ and } p_{\text{false}}(i, \beta, T) = 0, \tag{4.2}$$

where $T$ was set to the smallest segment length $T_1$, was searched. If such values existed for all four frequencies, the optimal time window was set to $T_1$ and the optimal thresholds $\beta_i$ were set to the corresponding value of $\beta$ satisfying (4.2). If such $\beta$ did not exist for at least one of the four stimulation frequencies, $T$ was set to the next higher segment length, and again, thresholds satisfying the conditions (4.2) were searched. This procedure was repeated with increasing time window lengths until suitable thresholds satisfying (4.2) were found. Table 4.3 shows an example of the threshold calibration process.

### Three-step spelling application

The three-step spelling application presented in section 4.3 was used to test the on-line performance after calibration. Two small changes were made: The language of the speller was set to English, and the size of the boxes was decreased to 125 × 125 pixels. Again, a sliding classification time window was used [63], with the modification that the minimal time window was determined by the wizard software.

## Results

This section presents an overview of the wizard outputs, the results of the on-line spelling performance, and the questionnaire answers.

### Wizard

In total, 70 s of recorded EEG data were analyzed during the three phases of the wizard (10 s for the first, 20 s for the second, and 40 s for the third phase).

The second phase (multi-target stimulation) yielded a set of four optimal stimulation frequencies. Figure 4.7 A shows the percentage of participants each candidate frequency was determined for by the wizard. The most frequently determined stimulation frequency was 7.5 Hz. Generally, frequencies lower than 8 Hz were suggested more often than higher frequencies. The highest candidate frequencies 13.33, 15, 17.14, and 20 Hz were not determined at all.

In the third phase (sequential stimulation), the selected stimuli were presented in sequence, and an optimal minimal time window was determined. Figure 4.7 B shows the percentage of participants each minimal time window was determined for. For most participants, the smallest time window, 813 ms, was selected; the longest time window, 8125 ms, was not selected for any user.

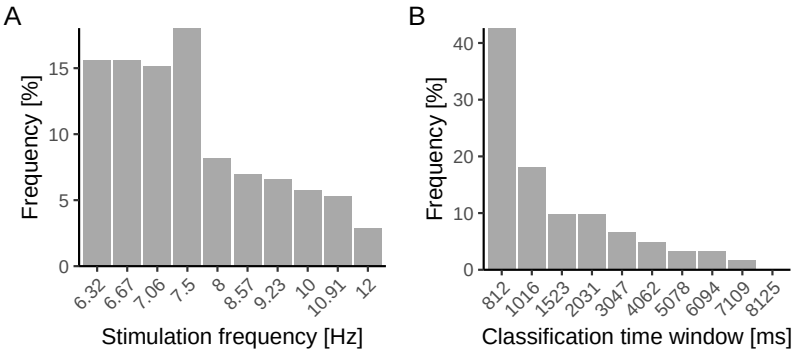



**Figure 4.7:** Outputs of the wizard software. (A) Distribution of stimulation frequencies determined by the wizard as part of the optimal set. (B) Relative frequencies of optimal classification time windows determined by the wizard. Data from Gembler et al. [86].

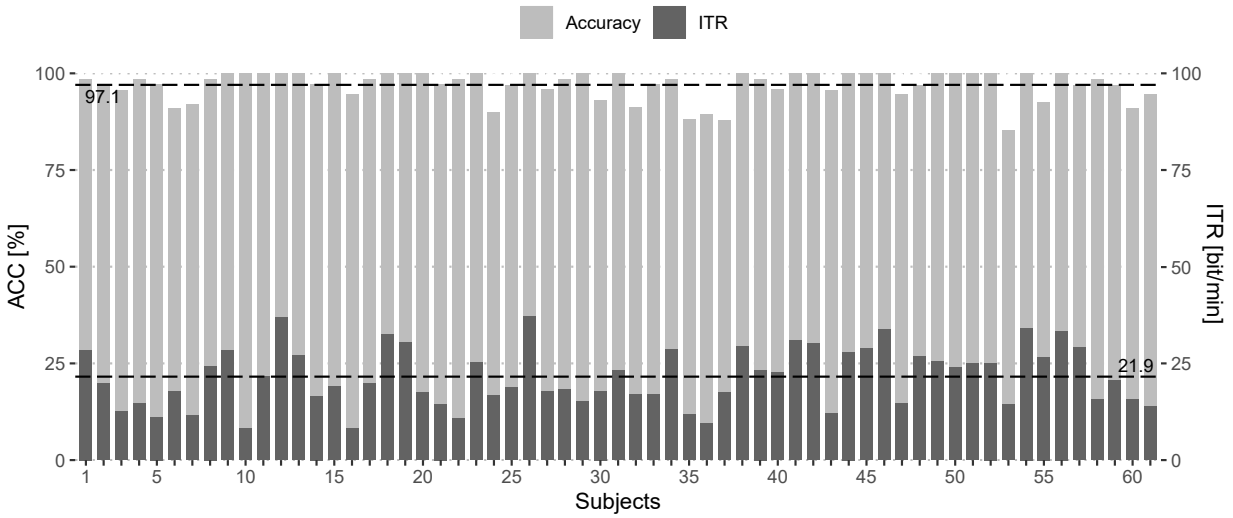



**Figure 4.8:** Individual ITRs and accuracies of the on-line experiment. The dashed lines indicate the mean values. Data from Gembler et al., 2015 [86].

### On-line spelling performance

All 61 subjects were able to complete the spelling task; None of the subjects reported any discomfort during the experiment. Figure 4.8 shows the overall results of the spelling task. The analysis of the spelling performance reveals a mean (SD) ITR of 21.9 (7.6) bpm, a mean accuracy of 97.1 (3.7)%, and mean values for OCM of 3.7 (1.2) char/min. All participants reached accuracies above 85%; 24 of the 61 subjects even completed the spelling task without errors, achieving an accuracy of 100%.

### Impact of gender on BCI performance and questionnaire results

Figure 4.9 shows the differences in BCI performance for female and male participants. Female participants reached an ITR of 25.4 (6.5) bpm and males reached an ITR of 20.1 (7.3) bpm; a Welch's $t$-test revealed a significant difference between these mean values ($t = 2.64$, $p = 0.012$). Similarly, significant differences in accuracy and OCM were found according to Welch's $t$-tests: The mean (SD) accuracies were 98.5 (2.6)% and 96.5 (4.0)% for female and male participants; the difference of means was significant ($t = 2.34$, $p = 0.024$). The mean (SD) values for OCM were 4.3 (1.1) char/min and 3.5 (1.2) char/min for female and male participants; the difference of means was significant ($t = 2.46$, $p = 0.019$).

In the pre- and post questionnaires, participants were asked questions regarding gender, the need for vision correction, tiredness, and BCI experience. For most subjects, the calibration and the use of the spelling interface did not produce fatigue, 5 participants

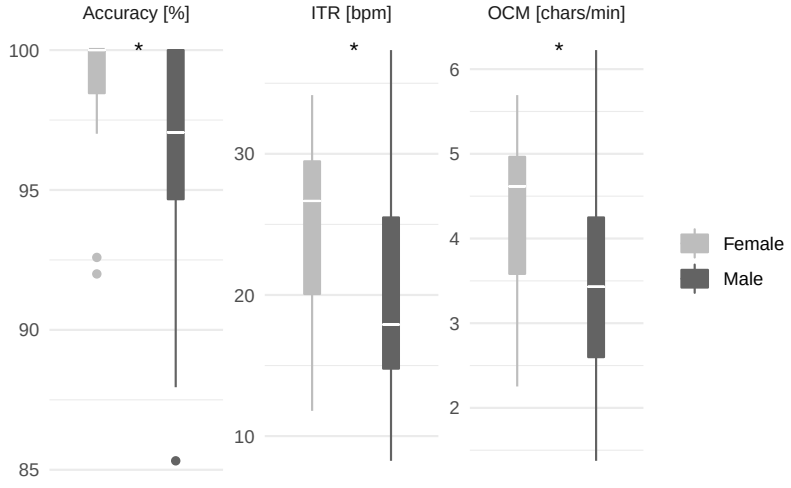



**Figure 4.9:** Comparison of SSVEP on-line spelling performances of female and male participants. The task was to spell the sentence 'RHINE WAAL UNIVERSITY'. Compared are ITR, classification accuracy and output characters per minute (OCM). The significance of Welch's $t$-tests are marked by asterisks; * indicates $p < 0.05$. Data from Gembler et al. [86].

reported decreased tiredness, and 14 participants reported slightly increased tiredness. The questionnaire answers are provided in Table 4.4.

## Discussion

The presented wizard for SSVEP-based BCIs determined SSVEP key parameters in a short time. The calibrated system was tested with a four target on-line spelling application and yielded high classification accuracies and a literacy rate of 100%.

Only a single click was required for the calibration procedure. Thus inexperienced personnel could set up the SSVEP system. Overall, including pauses between steps, the presented calibration process took less than two minutes.

Regarding the frequency choice, for the majority of subjects, lower frequencies were determined (see Figure 4.7 A). For nine participants, the lowest possible frequency set (6.32, 6.67, 7.06, and 7.50 Hz) was determined. Interestingly, the most frequently selected frequency was 7.5 Hz, which might be explained by the fact that its second harmonic is 15 Hz, which is the stimulation frequency at which the SSVEP response is maximum, according to Pastor et al. [144].

It should be noted that the presented wizard determined the optimal frequencies on the basis of the classification accuracy; user-friendliness was not taken into account. In general, users tend to be more annoyed by lower frequencies, and visual fatigue is more likely to occur (as discussed in section 4.1). Indeed, according to the user questionnaire, nearly a third of the subjects reported being annoyed by the flickering.

**Table 4.4:** Questionnaire results. The results are provided as number of respondents. Data from Gembler et al. [86].

| **Pre-questionnaire** | | **Number of respondents** | **Mean** | **SD** | **Range** |
|---|---|---|---|---|---|
| Age (Years) | | 61 | 22.8 | 5.0 | 17-49 |
| Gender | F | 44 | | | |
| | M | 17 | | | |
| Need for vision correction | Yes | 22 | | | |
| | No | 39 | | | |
| Hours of sleep last night | | | 6.8 | 1.2 | 4-9 |
| Are you tired? | | | 2.0 | 0.9 | 1-4 |
| (1) Not tired | | 19 | | | |
| (2) Little tired | | 25 | | | |
| (3) Moderately tired | | 15 | | | |
| (4) Tired | | 3 | | | |
| (5) Very tired | | 0 | | | |
| Did you use a BCI before? | Yes | 6 | | | |
| | No | 55 | | | |
| **Post-questionnaire** | | | | | |
| Are you tired? | | | 2.2 | 1.0 | 1-4 |
| (1) Not tired | | 17 | | | |
| (2) Little tired | | 20 | | | |
| (3) Moderately tired | | 20 | | | |
| (4) Tired | | 1 | | | |
| (5) Very tired | | 3 | | | |
| Was the flickering annoying? | Yes | 20 | | | |
| | No | 41 | | | |
| Would you recommend the BCI system? | Yes | 58 | | | |
| | No | 3 | | | |

A major challenge in SSVEP-based BCIs lies in finding a compromise between accuracy and speed. In asynchronous applications, such as the presented three-step spelling application, commands corresponding to the stimulation frequencies are produced only if their probabilities exceed predefined thresholds. The careful calibration of the thresholds and the minimal time windows balances system speed and system accuracy.

Figure 4.7 B reveals a wide variety of ideal minimum time windows among the 61 participants; The longest minimal time window determined was 7109 ms, but in most cases, the shortest possible window length of 812 ms was determined. Although longer time windows result in lower ITR, they yield higher accuracies, and for some users, they are necessary to achieve reliable control. Thus, user-specific calibration, as provided by the presented wizard, is a necessary step to ensure a high BCI literacy rate.

All participants, 100%, achieved more than 85% accuracy, 93% achieved more than 90% accuracy, 77% achieved more than 95% accuracy, and still, 39% achieved 100% accuracy. These accuracies are comparably high; in a previous smaller sized study, the same

**Table 4.5:** Comparison of SSVEP-BCI results of several larger sized studies. In the first two studies (2009 and 2011) all BCI illiterate subjects were excluded from further calculation of mean values. Table modified from Gembler et al. [86].

| | Volosyak et al. 2009 [1] | Volosyak et al. 2011 [8] | Guger et al. 2012 [142] | Gembler et al. 2015 [86] |
|---|---|---|---|---|
| Number of subjects | 37 | 86 | 57 | 61 |
| Mean accuracy [%] | 92.9 | 92.3 | 95.5 | 97.1 |
| Literacy rate [%] | 86.5 | 97.7 | 100 | 100 |
| Number of classes | 5 | 4 | 4 | 4 |
| Time-window [s] | 2 | 2 | 3 | 0.8 - 8 |

GUI as in the presented study was tested with six healthy subjects, and a mean accuracy of 87.4 (6.7)% was reached [130]; frequencies were not calibrated user dependently and time windows were not determined automatically. In comparison, the mean (SD) accuracy of 97.1 (3.7)% achieved in the presented study is significantly higher, which supports our hypothesis that the classification accuracy can be improved through automated user-specific parameter selection.

Closely related to the achieved classification accuracy is the BCI illiteracy rate. Throughout previous SSVEP studies, the BCI literacy rate has been gradually improved. Volosyak et al. [1] reported a BCI illiteracy rate of 14%. Due to further modifications, the BCI illiteracy rate was reduced to 2.33% two years later [8]. Guger et al. [142] showed that their SSVEP-BCI could provide effective communication for all 53 subjects. One explanation for a high literacy rate in SSVEP-systems is a low number of stimulation targets. Guger et al. [142] and Volosyak et al. [8] used only four stimulation frequencies (see Table 4.5). In addition, Guger et al. [142] used a relatively large classification time window of 3 s and achieved a BCI literacy rate of 100% (see Table 4.5).

It should be noted that BCI literacy among all participants was also achieved in studies using other BCI approaches. Kaufmann et al. [75] reported that all 19 subjects were able to complete a spelling task with a P300 speller with an average accuracy of 91.2% and an ITR of 15.1 bpm and in a study with 99 subjects, Guger et al. [141] reported a BCI literacy rate of 100% as well. Guger et al. [37] also achieved full BCI literacy with 81 subjects using the motor imagery paradigm.

Interestingly, in the presented study, a significant difference in BCI performance between female and male participants was observed. Trends that female subjects performed better were also observed in previous works [8, 82]. Further investigations are needed to establish to what extent the difference is dependent on the frequency range. As investigated in the previous section, subject age impacts performance as well. In this sense, the presented study might not

be reflective of the general population due to the low mean age of 22.8 years.

It should be noted that on rare occasions (four participants in total), the determined classification threshold corresponding to the fourth target (box 'Del') was too low, resulting in poor performance during the familiarization run. In these cases, the calibration process was repeated. An explanation for this issue is that the wizard did not consider the spatial arrangement of the boxes. Especially the undo function, i.e., the box containing 'Del' or 'Back', had a rather prominent position (see Figure 4.2). Further software improvements are necessary; the wizard may be integrated directly in the applications, such as the three-step speller so that spatial proximity and the size of the targets in the calibration GUI and the spelling GUI are consistent.

In summary, the study demonstrates that through user-specific parameter setup, reliable BCI control can be achieved by a broad population. The presented system yielded a 100% literacy rate and high accuracies for the following reasons:

- frequencies, time windows, and thresholds were calibrated individually for each user;
- sufficiently long classification time windows were used for poor performers; and
- the number of stimulation frequencies was only four.

The latter point will be addressed more thoroughly in the next section, where we investigate the impact of the number of SSVEP targets on performance. In principle, the presented calibration methods are extensible to systems with a higher number of targets.

## 4.5 Impact of the Number of Targets on SSVEP Performance

This section is an amended version of [61]: Gembler et al. (2017), 'Suitable Number of Visual Stimuli for SSVEP-Based BCI Spelling Applications', [116]: Gembler et al. (2016), 'Exploring the Possibilities and Limitations of Multitarget SSVEP-Based BCI Applications'.

Although the mean accuracies achieved with the four-target spelling application were quite high (see, section 4.3 and section 4.4), a higher ITR is desirable.

One way to achieve higher ITRs is by increasing the number of targets. It is worth investigating to what extent the accuracy and BCI literacy depend on this variable. Two studies were conducted to investigate the impact of the number of targets on BCI literacy rate, classification accuracy, and ITR.

In the first study, 'Exploring the Possibilities and Limitations of Multitarget SSVEP-Based BCI Applications' [116], a cue guided on-line test was performed with different sized stimulus matrices.

In the second study, 'Suitable Number of Visual Stimuli for SSVEP-Based BCI Spelling Applications' [61] results from copy spelling tasks with 4, 6, and 28-target BCI spellers were analyzed. This section summarizes the two studies.

The number of classes of a BCI impacts the design of the GUI and is linked to its complexity. SSVEP-based spelling systems using stable frequencies by employing divisors of the monitor refresh rate, as discussed in section 3.1, typically present a single-digit number of targets. For example, the three-step speller presented in section 4.3 presents only four targets, the Bremen BCI presents five targets [42], the modified Bremen BCI with dictionary integration presents up to seven targets [76]. A significant disadvantage of these applications is that several steps are required to select a letter; the character output speed is therefore limited (see section 2.6).

Using the frequency approximation method [65], as described in section 3.1, more complex multi-target interfaces can be realized. Indeed, applications with twelve stimuli to select numbers to dial a telephone [145], and even 40 stimuli to select characters [52] have been developed.

Although multi-target BCIs allow high ITRs, a negative correlation between the number of targets and the accuracy has been repeatedly reported [103, 130, 146]. For example, Carvalho et al. [146] tested GUIs using different numbers of targets with two stroke patients and eight healthy participants. They reported mean accuracies of 97%, 77%, and 57% for a two, four, and six class interface, respectively.

The literacy rate is usually high for BCIs employing a low number of SSVEP stimuli. As discussed in section 4.4, some BCI studies with a large number of participants (>50) reported that all participants were able to gain control over applications employing four targets [86, 142].

The goal of the experiments presented in this section was to explore how many targets can be reliably distinguished using the SSVEP paradigm.

To this purpose, in [116], seven healthy subjects underwent a cue guided simulated on-line experiment with a variable number of targets; accuracy and ITR were measured for systems with 15, 24, 28, 35, 60, and 84 stimulation frequencies.

We also investigated the number of BCI targets under more practical conditions, as results from on-line and simulated on-line tasks usually differ. In [61], a series of copy spelling experiments were conducted to explore the suitable numbers of visual stimuli in terms of user-friendliness and performance. Three different custom-made

spellers were tested: a three-step speller with four stimuli, a two-step speller with six stimuli, and a single-step speller (resembling a German QWERTZ-style keyboard) with 28 stimuli.

## Methods

In this section, we summarize the methods and materials for the cue guided selection experiment [116]) and the copy spelling experiment [61], in the following referred to as matrix experiment and speller experiment. In terms of hardware, the setup presented in section 4.3 was used.

### Participants

In the matrix experiment, seven subjects (three females, four males) without disabilities participated. The mean (SD) age was 24.9 (3.8) years (range 22-30).

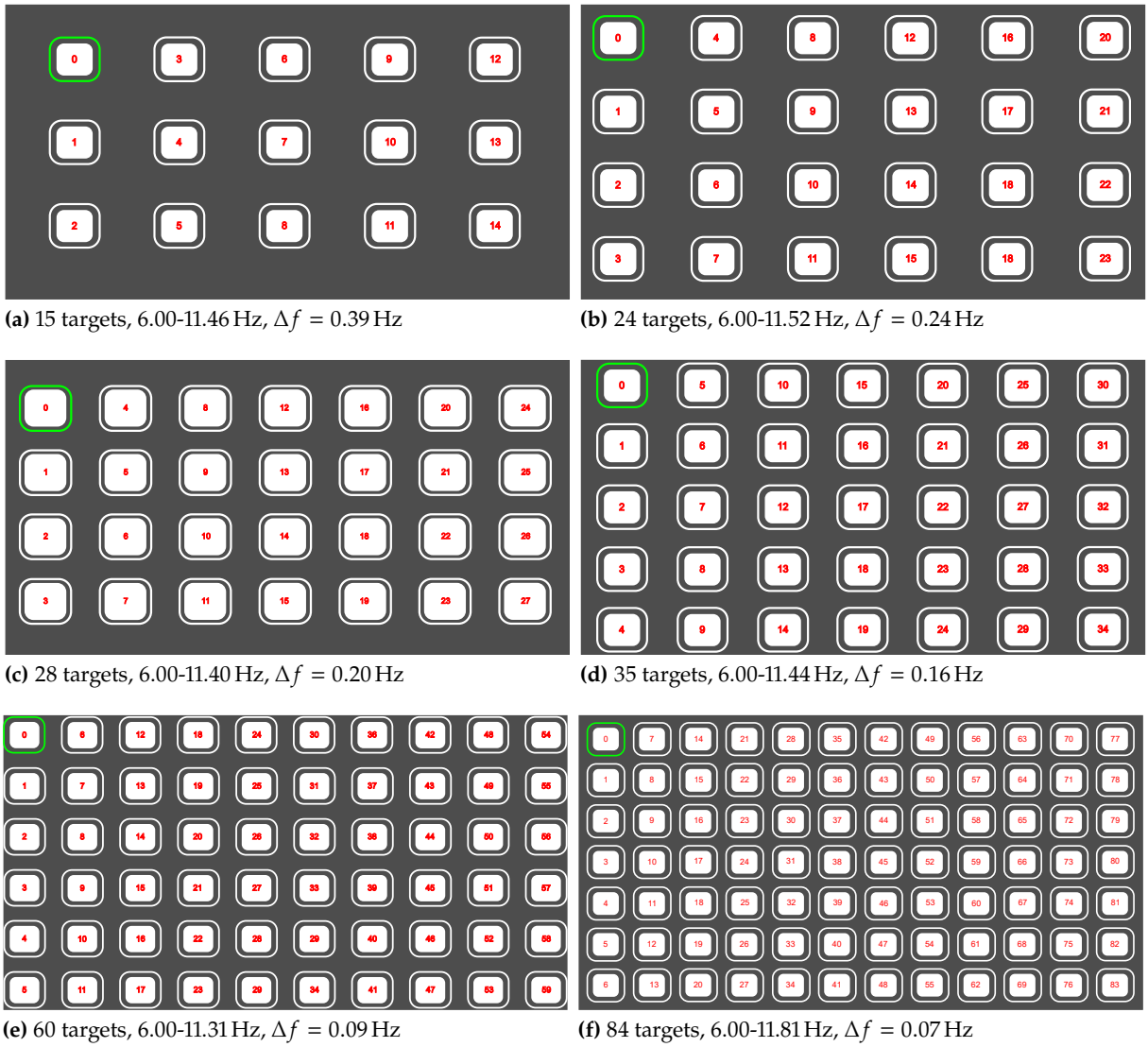
In the speller experiment, ten subjects (one female, nine males) without disabilities participated. The mean (SD) age was 25.5 (4.0) years (range 21-32).

All subjects had normal or corrected-to-normal vision; spectacles were worn if needed. All participants were students or employees of Rhine-Waal University. The subjects did not receive monetary compensation for participation.

### Experimental protocol

Participants performed cue guided selection tasks with six different stimulation matrices in the matrix experiment and copy spelling tasks with three different applications in the speller experiment. Figure 4.10 and Figure 4.11 show the GUIs used in the two experiments.

**Matrix experiment** Cue guided copy selection tasks with matrices of different sizes were performed (see Figure 4.10). Initially, a stimulus matrix presenting fifteen boxes, containing the numbers 0 to 14, was presented to the participant. A green frame highlighted the target the user needed to gaze at. When the target was classified correctly, the flickering paused for 1 s, and another box was highlighted. The boxes were highlighted in random order. When the participant successfully selected each box once, the next larger matrix was presented, the participant took a short break, and the procedure was repeated. If the participant could not complete the task with one of the stimulation matrices, the experiment ended;



**(a)** 15 targets, 6.00-11.46 Hz, $\Delta f = 0.39$ Hz

**(b)** 24 targets, 6.00-11.52 Hz, $\Delta f = 0.24$ Hz

**(c)** 28 targets, 6.00-11.40 Hz, $\Delta f = 0.20$ Hz

**(d)** 35 targets, 6.00-11.44 Hz, $\Delta f = 0.16$ Hz

**(e)** 60 targets, 6.00-11.31 Hz, $\Delta f = 0.09$ Hz

**(f)** 84 targets, 6.00-11.81 Hz, $\Delta f = 0.07$ Hz

**Figure 4.10:** Multi-target test matrices. Displayed are the GUIs of the test applications from [116]. (a-f) Provided are the number of classes, the frequency range, and the frequency resolution, i.e., the interval between neighboring stimuli, $\Delta f$. Participants performed cue guided selection tasks starting with the 15-target interface. The targets were cued in random order. When the selection task was completed, the next larger matrix was presented.

in this case, the tasks with a higher number of targets were not performed. For all stimulation matrices, equidistant stimulation frequencies were employed. The frequency resolutions ranged from 0.07 Hz for the 84-target stimulus matrix to 0.49 for the 15-target stimulus matrix., and the lowest frequency was 6 Hz for all matrices (see Figure 4.10).

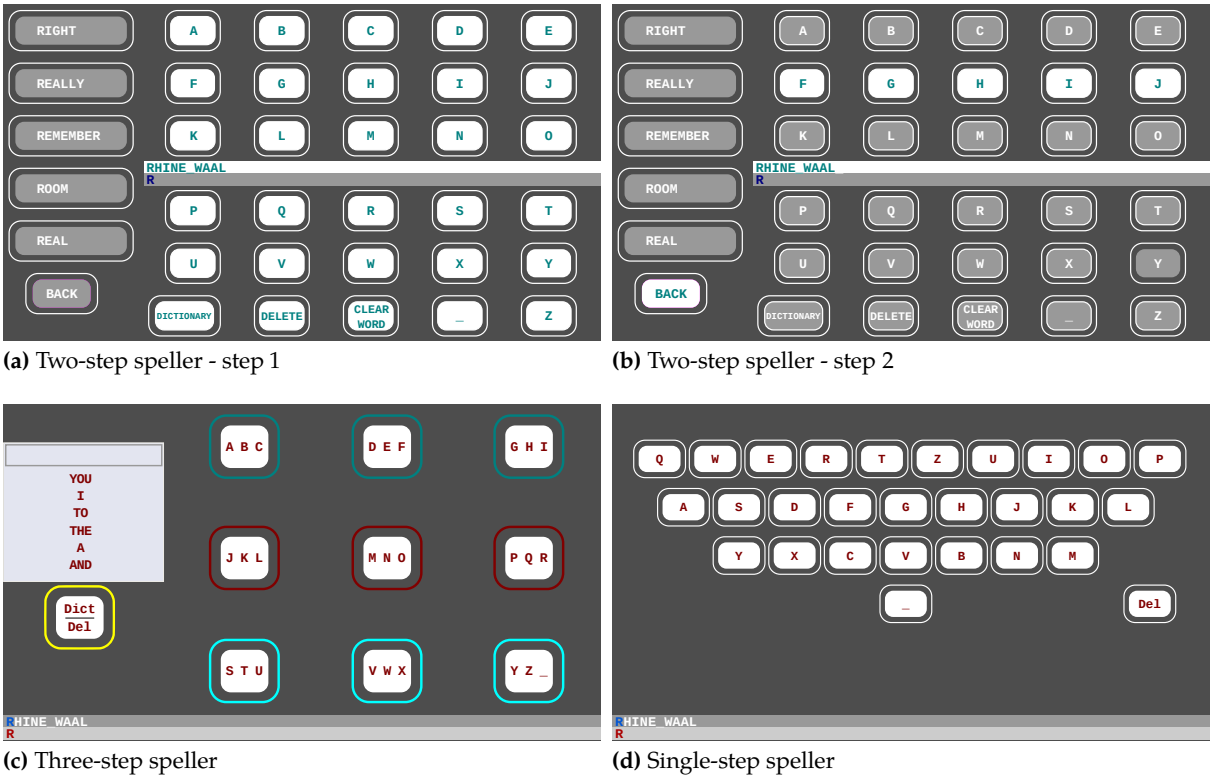
For signal classification, the MEC was used (see section 3.3), and for stimulus presentation, the frequency approximation method was used (see section 3.1).

**Speller experiment** Copy spelling tasks with three spelling applications were performed. The three-step speller and the two-step

speller allowed the selection of single letters and complete words in three or two steps, respectively. The single-step speller allowed the selection of letters in a single step. For all spellers, stimuli were represented by flickering boxes containing letters. To increase the user-friendliness of the spellers, we implemented an audio feedback corresponding to the generated outputs.

The GUI of the three-step speller was a modified version of the application presented in section 4.3. The four stimulation frequencies were selected as divisors of the vertical refresh rate, yielding a constant number of frames in each cycle, as explained in section 3.1. Before testing this application, participants went through the steps of the wizard software, which determined the stimulation frequencies (see section 4.4). In the first step, nine boxes, each containing three letters of the alphabet (26 letters plus and 'space'), were presented. The frames of the boxes had different colors, and each stimulation frequency was associated with a unique color (see Figure 4.11 c); an additional 10-th box, 'Dict/Del' (delete the last spelled character or switch to the dictionary mode) was displayed on the far left side of the screen. The sizes of the boxes varied in relation to the SSVEP power distribution between $140 \times 130$ and $230 \times 210$ pixels. In the second step, the boxes of the selected row were outlined with distinct colors (green, red, and blue), and the boxes of the remaining rows were grayed out. Only the frequencies (and the corresponding frame colors) changed; the position of the target letter remained the same, i.e., no gaze shifting was necessary between the first and the second step. The user was able to select one out of three boxes containing three letters each. In the third step, the boxes were rearranged. Three individual boxes, each containing individual letters, were presented. The functioning of the yellow-framed target (on the left) was dependent on the current selection step. In the first step, it enabled the user to delete the last selected letter or select word suggestions that were positioned above the 'Dict/Del' button. In the second and third steps, the far left box contained the command 'back', which lead to the previous step.

For the two-step speller, six frequencies were selected as divisors of the vertical refresh rate, which were also determined by the wizard software before the copy spelling task was conducted. A matrix presenting 30 white boxes (containing individual letters) and six additional gray boxes (containing five dictionary suggestions and a 'back' button) was displayed (see Figure 4.11 a). The sizes of the white boxes varied between $140 \times 90$ and $170 \times 130$ pixels, and the size of the gray boxes varied between $370 \times 90$ and $400 \times 130$ pixels in relation to the SSVEP power distribution. The lower row of the matrix contained the special targets 'Dictionary', 'Delete', and 'Clear word'. Selecting the 'Dictionary' button allowed the user to




**(a)** Two-step speller - step 1

**(b)** Two-step speller - step 2

**(c)** Three-step speller

**(d)** Single-step speller

**Figure 4.11:** Single-step, two-step, and three-step speller from [61]. (a) The two-step speller. The subject spelled the letter 'H'. Every row of the stimulation matrix flickered with a different frequency (same frequencies for columns). To select a letter, the user needed to select the corresponding row first. (b) In the second step, all characters of the selected row flickered with different frequencies; the desired letter 'H' could be selected. Characters could be selected in two steps; 6 frequencies were used. (c) The three-step speller. Characters could be selected in three steps; 4 frequencies were used. (d) The single-step speller. Characters could be selected in a single step; 28 frequencies were used.

choose one of up to five word suggestions. The boxes 'Delete' and 'Clear word' allowed the user to delete the last selected character or word, respectively. As only six distinct frequencies were used, each row of the white boxes was coded with a single stimulation frequency. For example, the boxes 'F','G','H','I', and 'J' all flickered in unison; by gazing at a particular box, the entire row was selected. In the second step (see Figure 4.11 b), the boxes of the selected row were coded with individual frequencies and the remaining rows were grayed out. Only six boxes (the five boxes from the selected row and the box containing 'back') were flickering. The output of the two-step speller was displayed at the center of the screen.

For the single-step speller, 28 frequencies were employed using the frequency approximation method, as explained in section 3.1. Frequencies ranging from 6.1 to 11.7 Hz (resolution 0.2 Hz) were used. The GUI displayed 28 buttons, which were arranged into four rows resembling a German QWERTZ-keyboard layout (see Figure 4.11 d). Each button flickered with a specific frequency. The size of the boxes varied between $130 \times 90$ and $170 \times 120$ pixels. The rows contained 10, 9, 7, and 2 buttons from top to bottom. A 'space' and a 'delete' button were presented in the bottom row.

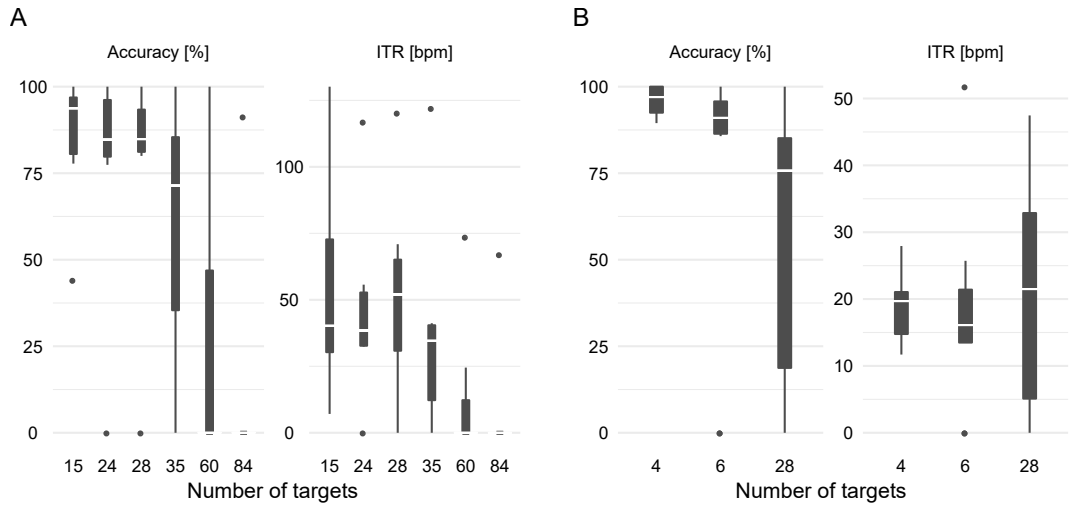



**Figure 4.12:** Impact of the number of stimuli on SSVEP-BCI performance. (A) Shown are the results from a cue guided on-line experiment utilizing 15, 24, 28, 35, 60, and 84-target matrices. (B) Shown are the results from on-line copy spelling experiments with the three-step, two-step, and single-step speller, using 4, 6, and 28 frequencies, respectively. The performance is compared on the basis of classification accuracy, ITR. In case a participant was unable to control a system, accuracy and ITR were set to 0. Data from Gembler et al. [116]. and Gembler et al. [61]

Subjects tested the spelling applications in random order. Prior to the copy spelling tasks, familiarization runs were performed. For all applications, the copy spelling task was 'RHINE WAAL'. Spelling errors needed to be corrected with the integrated delete function. In case a participant was not able to complete the task in a certain time frame, or if repeated misclassifications occurred, the task was stopped, and the next application was tested. The experiment took, on average, 40 minutes for each participant.

## Results

Figure 4.12 summarizes ITRs and accuracies for both experiments. In the following, the detailed results are provided.

**Matrix experiment** The performance was evaluated using classification accuracy and ITR. For each stimulus matrix, the individual accuracies and literacy rates are listed in Table 4.6 (a); the corresponding ITRs are listed in Table 4.6 (b).

To keep the duration of the experiment short, we decided to start the selection tasks with the smallest matrix and continued in ascending order of the number of frequencies used. If the subject could not control a system, the experiment ended. Therefore, results might be distorted by mental fatigue or training effect. To avoid a performance decrease due to fatigue, participants were instructed to take a short break before they continued with the next matrix.

**Table 4.6:** Accuracies, literacy rates, and ITRs achieved with the multi-target SSVEP test matrices. Seven healthy participants controlled stimulation matrices of different sizes. The dash indicates that no control over the system was achieved. Data from Gembler et al. [116].

**(a)** Accuracies and literacy rates

| **Subject** | **15 Targets** | **24 Targets** | **28 Targets** | **35 Targets** | **60 Targets** | **84 Targets** |
|---|---|---|---|---|---|---|
| **#** | Accuracy [%] | | | | | |
| 1 | 83 | 86 | 100 | 73 | - | - |
| 2 | 100 | 100 | 100 | 100 | 100 | 91 |
| 3 | 94 | 77 | 80 | - | - | - |
| 4 | 94 | 100 | 85 | 83 | 94 | - |
| 5 | 100 | 86 | 90 | 71 | - | - |
| 6 | 44 | - | - | - | - | - |
| 7 | 83 | 92 | 82 | 88 | - | - |
| **Literacy rate [%]** | 86 | 86 | 86 | 71 | 29 | 14 |

**(b)** Information transfer rate (ITR)

| **Subject** | **15 Targets** | **24 Targets** | **28 Targets** | **35 Targets** | **60 Targets** | **84 Targets** |
|---|---|---|---|---|---|---|
| **#** | ITR [bpm] | | | | | |
| 1 | 31.0 | 32.4 | 39.4 | 24.5 | - | - |
| 2 | 130.2 | 116.9 | 120.3 | 122.1 | 73.7 | 67.0 |
| 3 | 29.7 | 33.1 | 22.4 | - | - | - |
| 4 | 40.3 | 55.7 | 52.0 | 34.6 | 24.5 | - |
| 5 | 83.6 | 49.8 | 59.3 | 39.7 | - | - |
| 6 | 7.1 | - | - | - | - | - |
| 7 | 61.8 | 38.5 | 70.9 | 41.2 | - | - |

The number of targets yielding best performance varied strongly between participants. One participant (S6) had poor performance from the beginning, achieving an accuracy of only 44% for the 15-target system. For the remaining participants (six out of seven), reliable control was achieved up to the 28-target matrix. Two participants (S2 and S4) were able to complete the task with the 60-target matrix. Remarkably, one participant (S2) could still control the 84-target system (achieving 91.3% accuracy). For this matrix, the frequency resolution was only 0.07 Hz.

The peak ITR in the experiment was 130.15 bpm; it was achieved by S2 with 15 targets. Three participants (S1, S5, and S7) achieved their peak performance with 28 targets.

**Speller experiment** Figure 4.13 provides individual results of participants for the tested spelling applications; Table 4.7 summarizes these results. Every participant stated to be familiar with the QWERTZ-layout and reported to use computers regularly in daily life.

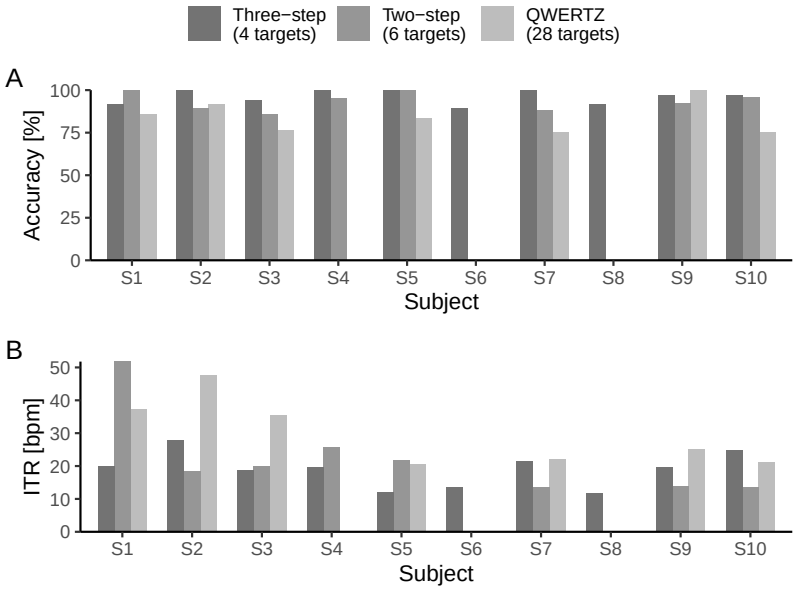



**Figure 4.13:** Impact of the number of stimuli on SSVEP-BCI spelling performance. (A) The classification accuracies and (B) the ITRs from the online copy spelling experiments of the three-step, two-step, and single-step speller using 4, 6, and 28 frequencies, respectively, are shown. If a participant was unable to control a system, accuracy and ITR were set to 0. Data from Gembler et al. [61]

All participants were able to control the three-step speller, reaching a mean accuracy of 96.1%. For the spellers with a higher number of targets, BCI literacy rate and average accuracy were lower. While 80% of the participants were able to control the two-step speller, only 70% were able to control the single-step speller.

The average command selection time windows (including gaze shifting phases) were 5.9, 6.6, and 7.3 s for the three-step speller, two-step speller, and single-step speller, respectively.

For the ITR calculation, the overall number of possible choices depended on the spelling application (4, 6, and 28 for the three-step speller, two-step speller, and the single-step speller, respectively). The highest ITR, 51.8 bpm, was achieved by S1 with the two-step speller. Four participants (S2, S3, S7, and S9) achieved their peak performance with the single-step speller.

**Table 4.7:** Performance comparison between three-step, two-step, and one-step SSVEP-BCIs. Participants that could not successfully control a speller were excluded from the calculation of mean values for that particular system. Data from Gembler et al. [61].

| | Number of displayed stimuli | Literacy rate [%] | Max ITR [bpm] | Mean ITR [bpm] | Mean ACC [%] | Mean time window [s] |
|---|---|---|---|---|---|---|
| **Three-step speller** | 4 | 100 | 27.9 | 18.9 | 96.1 | 5.9 |
| **Two-step speller** | 6 | 80 | 51.8 | 22.3 | 93.4 | 6.6 |
| **Single-step speller** | 28 | 70 | 47.5 | 29.2 | 83.9 | 7.3 |

### Discussion

Both experiments confirm that increasing the number of targets can decrease classification accuracy and BCI literacy rate. In both experiments, participants stated that it was harder to focus on the boxes with the higher target systems. Due to the closer proximity of targets, neighboring stimuli contributed to the SSVEP response and were sometimes falsely classified. A similar phenomenon has also been reported for P300-based systems (see, e.g., [147]).

Regarding the BCI speller experiment, performance varied considerably between participants. Some subjects achieved maximum ITR with the one-step speller (28 stimuli), others with the two-step speller (6 stimuli). One participant (S10) even reached the highest ITR with the three-step speller (4 stimuli), mainly due to lower accuracies with the other spellers.

The overall highest accuracies were achieved with the three-step speller, which was the only system that could be controlled reliably by all participants. For the higher target systems, the accuracies and literacy rates were lower. This is in line with other studies which reported a drop in accuracy when the number of stimuli increased (e.g., [146]).

Apart from the higher accuracy and literacy rate, a further advantage of the three-step speller is that it was least stressful for the participants. According to the questionnaire answers, the single-step speller was the most fatiguing application out of the three. Some participants also stated that they became frustrated due to the lower accuracy of the single-step system. In addition to that, the time windows for accurate classification of SSVEP responses were generally larger, if more stimuli were used.

As the optimal number of targets in terms of accuracy and ITR differed among participants, an additional calibration step to determine the optimal number of stimuli could be integrated into the wizard software. However, the aspect of user-friendliness needs to be taken into consideration as well; some users may prefer high accuracies and a less visual fatiguing interface over a high ITR.

# Investigating c-VEP parameters 5

The main difference between the c-VEP and SSVEP approach is that the former requires a training session where personal EEG templates are generated. Recent research strongly suggests that using personal EEG data rather than frequency templates yields higher accuracies (for a detailed review on this topic, see Zerafa et al. [114]). As BCI illiteracy is, in essence, defined by accuracy thresholds, personalized EEG templates may also reduce BCI illiteracy.

Due to the required synchronization between stimulus presentation and hardware acquisition, c-VEP systems are typically realized with fixed classification time windows (synchronous BCIs).

For synchronous systems, the manipulation of two key parameters, the time window and the classification thresholds, which has improved the BCI literacy rate for SSVEP systems, cannot be performed. The initial goal, therefore, was to develop an asynchronous c-VEP application.

The first study in this chapter introduces an eight-target c-VEP spelling application that achieves this goal (section 5.1). The presented speller incorporates novel features such as a sliding window mechanism and an $n$-gram based dictionary integration. The chapter further addresses the impact of user age and flickering speed on c-VEP performance (section 5.2). Following that, automatic calibration of ideal time windows and thresholds are explored. In addition to that, the stimulus paradigms SSVEP and c-VEP, both implemented using personal EEG data, are compared (section 5.3). Lastly, the sliding window mechanism and the $n$-gram based dictionary integration were adopted to a multi-target interface (section 5.4).

All sections presented in this chapter are based on published journal articles and conference proceedings about c-VEP-based BCIs [72, 148–150].

## 5.1 A Dictionary-driven Asynchronous c-VEP Spelling Application

Typical use cases of c-VEP-based BCIs are spelling applications for people with severe disabilities [13]. An issue with c-VEP-BCIs is that usually fixed time windows are used to produce command outputs.

This section is an amended version of [72]: Gembler and Volosyak (2019), 'A Novel Dictionary-Driven Mental Spelling Application Based on Code-Modulated Visual Evoked Potentials'.

Almost all c-VEP systems employ a fixed stopping approach, i.e., the classification is performed after exactly one full stimulation cycle has been completed. After that, the flickering usually stops giving the user a short time for gaze shifting. The length of the code pattern then becomes a bottleneck for the overall responsiveness of the system. Furthermore, classification methods based on fixed time windows do not take individual differences between users into account. As discussed in the previous chapter, the required time window to produce accurate commands can vary strongly between participants.

Thielen et al. [53] developed a 36-target VEP spelling application based on pseudo-random gold codes and implemented a criterion to decide when trials can be stopped. They reported an average accuracy of 86%, an ITR of 48 bpm, and an OCM of approximately 9 char/min. Still, an output was produced, when the classification window reached the trial length. It is, however, preferable that the system is also able to distinguish between intentional and unintentional target fixations. This way, more complex interfaces could be realized.

The study reported in this section presents a dynamic classification time window mechanism based on classification thresholds for the c-VEP paradigm. The classification confidence (certainty) was checked each calculation interval before the system produces outputs. This approach could reduce misclassifications and improve the overall usability of the system. It further allows the integration of prediction models (see section 2.6); if unintentional fixations do not produce output commands, the user has enough time to process changing elements of the user interface. Word prediction methods can enhance the character output speed, even if a low number of targets is used. In this study, an $n$-gram word prediction model was employed [89]. While prediction models based on $n$-gram models are quite common in BCIs [73], they are typically applied on the character level; in the presented study, the $n$-gram model was applied on the word level.

Regarding the signal classification, ensemble-based methods, which are usually used in machine learning, have recently boosted performance in SSVEP-based BCIs [52]. Here, an ensemble-based classification method was adopted for the spatial filters used with the c-VEP paradigm.

In summary, the contributions of the research were the following:

- improvements in c-VEP speller usability due to the dynamic sliding classification window mechanism;
- improvements in c-VEP detection due to an ensemble-based classification approach; and

- improvements in output effectiveness due to the integration of an $n$-gram based word prediction model.

The study evaluated the feasibility of the proposed methods based on an on-line experiment with 18 healthy participants. To this end, an eight-target two-step spelling interface with $n$-gram based word prediction feature providing three word suggestions was developed.

## Methods

This section describes the subject group, the hardware setup, the software, and the experimental design. The detailed descriptions of the sliding window mechanism and the dictionary-driven speller were also published in our publication [149].

### Participants

Eighteen non-disabled participants (eight female and ten male) with mean (SD) age of 23.3 (4.4) years, ranging from 19 to 31, were recruited from the Rhine-Waal University of Applied Sciences. Participants had normal or corrected-to-normal vision. In this study, participants received a small financial reward.

### Hardware

The used computer (MSI GT 73VR with Nvidia GTX1070 graphics card) operated on Microsoft Windows 10 Education running on an Intel processor (Intel Core i7, 2.70 GHz). A liquid crystal display screen (Asus ROG Swift PG258Q, 1920 × 1080 pixel, 240 Hz refresh rate) was used to display the user interface and to present the stimuli.

All 16 channels of the EEG amplifier (g.USBamp, Guger Technologies, Graz, Austria) were used; the electrodes were placed according to the 10-5 system of electrode placement (see section 2.2 for more details): Pz, P3, P4, P5, P6, PO3, PO4, PO7, PO8, POO1, POO2, O1, O2, Oz, O9, and O10. In general, reliable control may be achieved with a smaller number of channels. However, a high number of channels is needed to maximize accuracies and ITRs. The impact of the number of electrodes on accuracy is addressed in the results section. The common reference electrode was placed at Cz and the ground electrode at AFz. Standard abrasive electrolytic electrode gel was applied between the electrodes and the scalp to bring impedances below 5 kΩ. The sampling frequency of the amplifier, $F_s$, was set to 600 Hz.

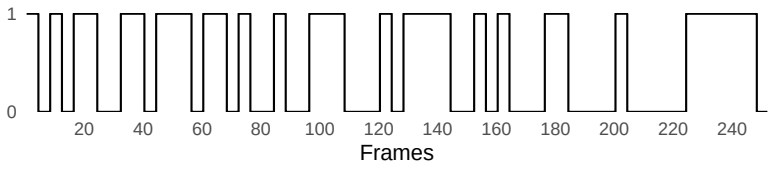



**Figure 5.1:** Stimulus pattern of the 63 bit $m$-sequence used in the experiment. The refresh rate of the monitor was 240 Hz. Each '1' corresponded to four frames where the associated stimulus was shown and each '0' to four frames where the stimulus was not shown. The duration of the stimulation cycle was 1.05 s. Figure adapted from Gembler and Volosyak [72].

### Stimulus design

In the c-VEP system used in this study, $K = 8$ stimulus classes were used. The color of a target stimuli alternated between 'black' (the color of the background, represented by '0') and 'white' (represented by '1') in accord with the underlying code pattern. To this end, the well-established 63 bit $m$-sequences were used (please refer to section 3.2 for more details).

The $m$-sequences $c_i$, $i = 1, \ldots, K$ were assigned to the stimulus matrix employing a circular shift of 2 bit ($c_1$ had no shift, $c_2$ was shifted by 2 bit to the left, $c_3$ was shifted by 4 bit to the left, etc.).

As explained in section 3.2, $m$-sequences can be generated with LFSRs. In all experiments in this chapter, the initial code $c_1$ was generated with an LFSR represented by the generator polynomial $x^6+x^5+1$ with the initial value 111110. Figure 5.1 shows the resulting code pattern.

In terms of synchronization between stimulus presentation and data acquisition, the software-based approach, described in section 3.6, was used.

### Experimental protocol

First, each of the 18 participants went through a training phase, which was required to generate individual templates and spatial filters for on-line classification. After that, on-line copy spelling tasks were performed.

In the training phase, data for each of the stimuli were collected. The data collection was grouped in six blocks, $n_b = 6$; in each block, each of the $K = 8$ targets was gazed at once. Hence, $n_b \cdot K = 48$ trials were collected in total.

Each of these trials lasted for 3.15 s, i.e., the code patterns repeated for 3 cycles. A green frame highlighted the box at which the user needed to focus. At the beginning of each of the $n_b$ recording blocks, the user started the flickering by pressing the space bar. After each trial, the next box the user needed to focus on was

**Table 5.1:** Individual sentence tasks of the on-line experiment.

| # | Sentence | # | Sentence |
|---|---|---|---|
| 1 | I FORGOT TO DO MY HOMEWORK | 10 | THE DIVING SUIT IS TOO SMALL |
| 2 | I LIKE TO EAT CHEESE | 11 | THE SUN IS SLOWLY RISING |
| 3 | I BOUGHT EGGS TODAY | 12 | IT IS GOING TO RAIN TOMORROW |
| 4 | I COULD NOT HEAR THAT | 13 | THE DOG BARKED LOUDLY |
| 5 | I DO NOT SPEAK FINNISH | 14 | THE LIGHT BULB HAS BURNED OUT |
| 6 | WHAT DID YOU HAVE IN MIND | 15 | HE SANG OUT OF TUNE |
| 7 | I AM NOT YET HUNGRY | 16 | MY BIKE HAS NOT BEEN STOLEN |
| 8 | HOW LATE IS IT | 17 | THEY OWN A BLACK CAT |
| 9 | I COULD EAT PIZZA EVERYDAY | 18 | AND THAT IS IT |

highlighted, and the flickering paused for one second. The boxes were highlighted sequentially from upper left to lower right. After every eighths trial (i.e., after every block), the user was allowed to rest.

The training phase was followed by a familiarization run where participants spelled the word BCI. The classification threshold for the certainty was adjusted manually during this familiarization run to ensure adequate speed and accuracy.

Three spelling tasks were performed: First, the word BRAIN was spelled (letter-by-letter task), thereafter, the sentence THAT_IS_-FUN was spelled (familiarization task, to learn the functioning of the integrated dictionary), and lastly, an additional sentence, different for each user was spelled (individual sentence task, see Table 5.1). For each task, participants needed to correct errors using the integrated UNDO function. In the sentence spelling tasks, dictionary suggestions could be selected.

#### Dictionary-driven spelling application

The dictionary-driven eight-target spelling interface presented eight boxes (230 × 230 pixel), each corresponding to one of the $K = 8$ stimulus classes, arranged as $2 \times 4$ stimulus matrix, were presented. The first row of the GUI contained 28 characters (26 letters, underscore, and full stop character) divided into four boxes (seven characters each). The second row offered three dictionary suggestions and a correction option. By selecting the correction option, the last typed character or word was deleted. By selecting a letter group from the first row, the associated characters were presented individually; thus, selecting individual characters required two steps.

The copy spelling sentence and the user output were presented in the center of the screen. If the classifier produced an output

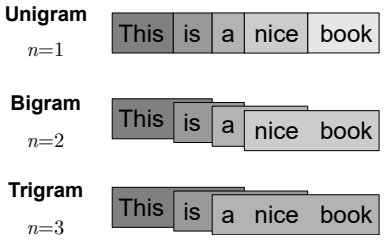



**Figure 5.2:** Item groupings of $n$-grams on the word level.

command, audio and visual feedback were provided: The size of the selected box increased for a short time, and a sound file, voicing the selected command, was played. Additionally, a progress bar displayed the current certainty level of the associated class label.

The dictionary suggestions were updated after each performed selection on the basis of an $n$-gram prediction model, which is used in computational linguistics. This model considers a sequence of $n$ items from a text database. For a given sequence of items (here words), $x_{i-(n-1)}, \ldots, x_{i-1}$, the next item $x_i$ is suggested on the basis of the conditional probabilities $P(x_i|x_{i-(n-1)}, \ldots, x_{i-1})$. In this application, an $n$-gram of size 2 (also called bi-gram, see Figure 5.2) was integrated, i.e., next word candidates were ranked according to their probability on the word level.

The text database was extracted from the Leipzig Corpora Collection, a ready to use corpora [151]. It contains a word frequency list and a list of word bi-grams (co-occurrences as next neighbors) containing observed frequency counts, which were generated from approximately 1 million sentences publicly accessible. Structured query language (SQL), a query language for relational databases, was used to retrieve word suggestions from the Leipzig text database. Based on the already typed part of the current word, three word suggestions were extracted using SQL statements. First, all co-occurrence pairs, including the previously typed word and the words beginning with the already typed part of the current word, were ordered according to their frequency. If this procedure resulted in less than three candidates, the suggestions were complemented with the word frequency list (independent of the precedent word), i.e., the most frequent words matching the already typed string were added. An example of the functioning of the dictionary-driven speller is provided in Figure 5.3.

### Spatial filtering and template matching

In this study, two approaches of spatial filtering, the conventional approach (as described in section 3.5) and an ensemble-based approach, were investigated. In both cases, the CCA (see section 3.4) was applied to design the spatial filters.

Each training trial was stored in an $m \times n$ matrix, where $m$ denotes the number of electrode channels (here all 16 signal channels of the amplifier were used for computation, i.e., $m = 16$) and $n$ denotes the number of sample points (here, $n = 1.05 \cdot F_s \cdot 3 = 1890$).

In the conventional template matching approach, as introduced in section 3.5, all training trials are circularly shifted to match the phase of the first trial. The shifted trials, denoted $\mathbf{Z}_i$, $i = 1, \ldots, n_b K$, were then averaged yielding an averaged template $\bar{\mathbf{Z}}$.


**Figure 5.3:** GUI of the dictionary-driven c-VEP spelling application. A participant was spelling the sentence "JUST_DO_IT". The selection of individual letters required two steps: First, the group containing the character needed to be selected (Layer I), and second, the box with the desired character needed to be selected (Layer II). In total, seven selections were required to complete the sentence (10 characters). Due to the sliding time window mechanism, there were no time constraints for selections.

With these trials, the matrices

$$\mathbf{Z} = [\mathbf{Z}_1\mathbf{Z}_2 \dots \mathbf{Z}_{n_bK}] \quad \text{and} \quad \mathbf{Z} = [\underbrace{\mathbf{Z}\mathbf{Z} \dots \mathbf{Z}}_{n_bK}] \tag{5.1}$$

were inserted into the CCA formula (section 3.4), yielding a filter vector $\mathbf{w}^{(1)} = \mathbf{w}_{\hat{\mathbf{Z}}}$. Class-specific templates $\mathbf{X}_i^{(1)}$, $i = 1, \dots, K$ were generated by circularly shifting $\mathbf{Z}$ in accordance with the bit-shift of the underlying code $c_i$ (please also refer to section 3.5 for more details).

For the ensemble-based approach, individual templates $\mathbf{X}_i^{(2)} \in \mathbb{R}^{m \times n}$ and filters $\mathbf{w}_i^{(2)}$ were determined for each stimulus ($i = 1, \dots, K$) independently. Let $\mathbf{T}_{i,j}$, $i = 1, \dots, K$, $j = 1, \dots, n_b$ denote the trial of class $i$ recorded in block $j$. Class-specific trial averages $\mathbf{X}_i$ were generated by averaging all trials corresponding to the $i$-th class,

$$\mathbf{X}_i = \frac{1}{n_b} \sum_{j=1}^{n_b} \mathbf{T}_{i,j}, \quad i = 1, \dots, K. \tag{5.2}$$

To design the spatial filters, for all classes $i = 1, \dots, K$ the matrices,

$$\mathbf{T}_i = [\mathbf{T}_{i,1}\mathbf{T}_{i,2} \dots \mathbf{T}_{i,n_b}] \quad \text{and} \quad \mathbf{X}_\mathbf{i}^{(2)} = [\underbrace{\mathbf{X}_i\mathbf{X}_i \dots \mathbf{X}_i}_{n_b}] \tag{5.3}$$

were constructed and, using CCA, filter vectors $\mathbf{w}_i^{(2)} = \mathbf{w}_{\hat{\mathbf{T}}_i}$, $i = 1, \dots, K$ were created.

For both methods, the on-line classification was performed after receiving new EEG data blocks, which were automatically added to a data buffer $\mathbf{Y} \in \mathbb{R}^{m \times n_y}$ with dynamically changing column dimension $n_y$.

The data buffer $\mathbf{Y}$ was compared to reference signals $\mathbf{R}_i^{(j)} \in \mathbb{R}^{m \times n_y}$, $i = 1, \dots, K$ which were constructed as sub-matrix of the corresponding training template from rows $1, \dots, m$ and columns $1, \dots, n_y$ from $\mathbf{X}_i^{(j)}$ for the conventional ($j = 1$) and ensemble method ($j = 2$), respectively.

To identify the class label of the data contained in the buffer $\mathbf{Y}$, the classifier determined the Pearson correlations between the spatially filtered reference signals and the spatially filtered EEG data in the buffer. For the conventional approach, correlations $\lambda_k^{(1)}$, were determined as

$$\lambda_k^{(1)} = \rho \left( \mathbf{Y}^T\mathbf{w}^{(1)}, {\mathbf{R}_k^{(1)}}^T\mathbf{w}^{(1)} \right), \quad k = 1, \dots, K; \tag{5.4}$$

for the ensemble-based approach, the ensemble correlations, $\lambda_k^{(2)}$,

were determined as

$$\lambda_k^{(2)} = \rho\left(\begin{bmatrix}\mathbf{Y}^T\mathbf{w}_1^{(2)}\\ \vdots \\ \mathbf{Y}^T\mathbf{w}_K^{(2)}\end{bmatrix}, \begin{bmatrix}\mathbf{R}_k^{(2)^T}\mathbf{w}_1^{(2)}\\ \vdots \\ \mathbf{R}_k^{(2)^T}\mathbf{w}_K^{(2)}\end{bmatrix}\right), \quad k = 1, \ldots, K. \tag{5.5}$$

In both cases, the classification output class label $C$ was set to

$$C = \underset{k=1,\ldots,K}{\arg\max}\, \lambda_k^{(j)}, \quad j = 1, 2. \tag{5.6}$$

**Sliding window mechanism for c-VEP signal classification**

The output command corresponding to a classified label was only performed if a threshold criterion was met. In this regard, sliding classification time windows of dynamic length were used [63], i.e., in the case where no classification could be made, a new classification was performed, after receiving the new EEG data.

The multichannel EEG signals that were about to be classified were stored in a matrix $\mathbf{Y} \in \mathbb{R}^{m \times n_y}$, where $n_y$ represents the length of the classification time window in samples. In practice, $n_y$ needs to be selected carefully. Too small time windows can lead to errors; on the other hand, if $n_y$ is too large, data unrelated to the desired target (e.g., due to gaze movements at the beginning of the time window) stays in the buffer and remains to be considered for classification, which can slow down performance. Therefore, $n_y$ needs to be restricted,

$$n_{y_{\min}} \leq n_y \leq n_{y_{\max}}. \tag{5.7}$$

The selection of the minimum time window, the lower bound in (5.7), is critical. Recall that the duration of one stimulus cycle was 1.05 s and the number of samples collected in this period, $n_c$, was $n_c = 630$. The upper bound, $n_{y_{\max}}$, was selected as a multiple of $n_c$.

In the on-line BCI, the time window extended incrementally. The amplifier transfers EEG data in blocks $\mathbf{A}_i \in \mathbb{R}^{m \times n_a}$, where $n_a$ denotes the number of samples per block. For the implementation of the sliding window mechanism, $n_a$ was selected as divider of the cycle length, i.e., $n_a | n_c$.

A further restriction to $n_a$ was given by the amplifier manufacturer. For the g.USBamp, the buffer needed to contain at least 20-30 ms of data. Here, $n_a$ was set to 30 samples (50 ms recordings with the sampling rate of 600 Hz).

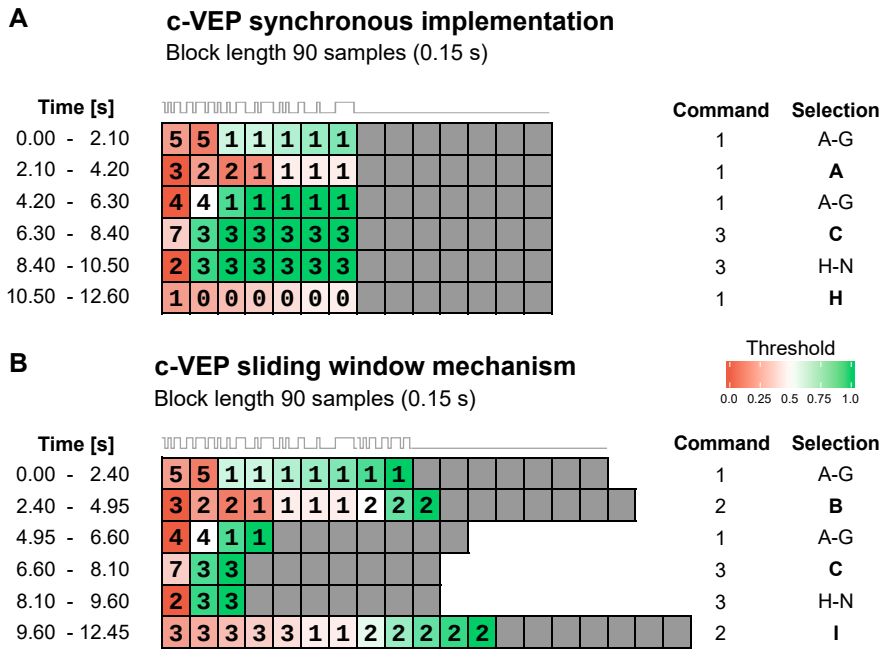



**Figure 5.4:** Illustration of the threshold-based sliding window mechanism used in the on-line experiment. Displayed are the classification times needed to spell the word BCI for (A) the synchronous implementation and (B) the threshold-based sliding window mechanism. Each row represents one classification. The cells contain the label classified after the received block and the certainty associated with the label (color-coded from red to green). The gray boxes indicate the gaze shifting phases (here, 7 blocks each). For the synchronous approach, output commands are is produced at fixed time intervals, i.e., after 1.05 s. For the sliding window mechanism, output commands are produced if a threshold criterion is met. In the example, the sliding window mechanism yields a shorter spelling time and higher accuracy. Figure adapted from Gembler and Volosyak [72].

The amplifier blocks were accumulated in a buffer $\mathbf{A} \in \mathbb{R}^{m \times n_{\hat{a}}}$,

$$\mathbf{A} = [\mathbf{A}_1 \mathbf{A}_2 \ldots].$$

If the number of samples of the buffer, $n_{\hat{a}}$, was too small, i.e., $n_{\hat{a}} < n_{y_{\min}}$, no classification was performed.

If $n_{y_{\min}} \leq n_{\hat{a}} \leq n_{y_{\max}}$, the classification time window gradually increased (with step width $n_a$). The classification was performed using the data matrix $\mathbf{Y} = \mathbf{A}$, i.e., all data from the buffer were considered for classification. If the classifier did not meet a threshold criterion, as described later, further EEG data were collected.

Lastly, if $n_{\hat{a}} > n_{y_{\max}}$, data were shuffled out. The data matrix $\mathbf{Y}$ was defined as the sub-matrix of $\mathbf{A}$ formed from rows $1, \ldots, m$ and columns $n_c k + 1, \ldots, n_{\hat{a}}$,

$$\mathbf{Y} = \mathbf{A}[1, \ldots, m \,;\, n_c k + 1, \ldots, n_{\hat{a}}],$$

where $k$ is the smallest integer such that $n_y$ is bound by (5.7). Since $n_a | n_c$, data collection and stimulus presentation remain synchronized.

The BCI output associated with a classified label $C$ was only

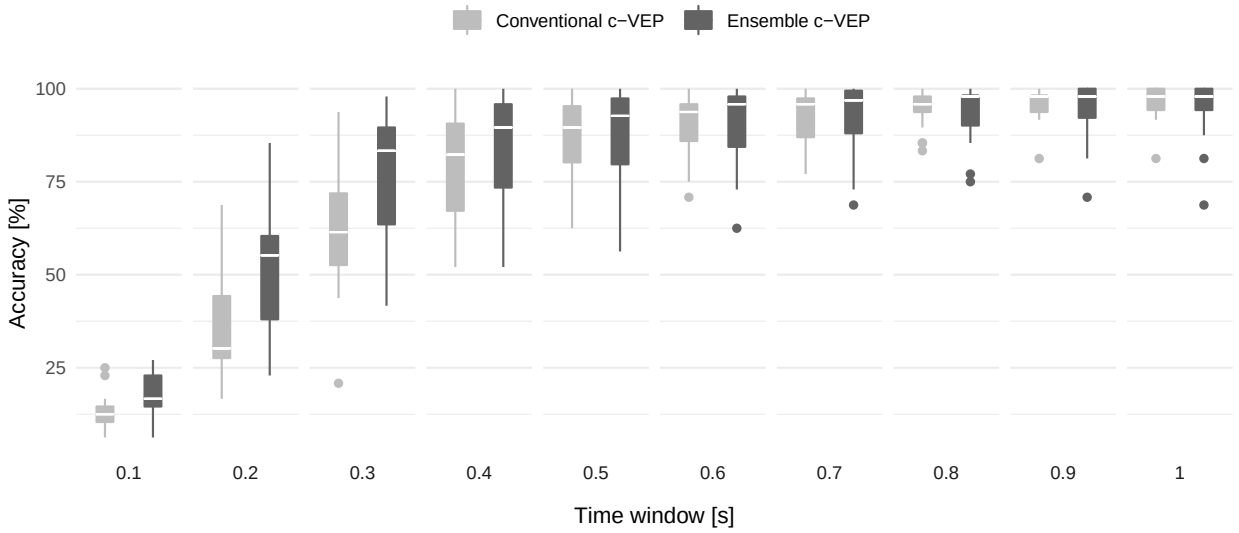



**Figure 5.5:** Classification accuracies achieved with the conventional c-VEP classification approach and the ensemble c-VEP classification approach. In the box plots, outliers (data points outside 1.5 times the interquartile range) are located outside the "whiskers".

performed if a threshold criterion was met, which is described in the following. The decision certainty, $\Delta_C$, which was determined as the distance between highest and second-highest correlation, needed to surpass a threshold value, $\beta$. This threshold was set for each participant individually after the training. In other words, the output command was only performed if $n_y \geq n_{y_{\min}}$ and $\Delta_C \geq \beta$.

After a produced output command, the data buffers $\mathbf{A}$ and $\mathbf{Y}$ were cleared, and a 2 s gaze shifting period followed. In this gaze shifting period, the amplifier data blocks, $\mathbf{A}_i$, were ignored and the stimuli did not flicker, allowing the user to shift his or her gaze to the next target. The BCI did not require a full cycle of the stimulation pattern for classification. If a command was classified before the stimulus pattern completed a full cycle, the flickering stopped. Figure 5.4 illustrates the sliding window mechanism and compares it to the conventional method.

## Results

The conventional and the ensemble-based classification approaches were compared off-line via stratified 6-fold cross-validation (see section 2.7). All but one recording blocks were used for the training, and one block was used as validation data. The cross-validation process was repeated $n_b$ times, with each recording block used once as the validation data. The $n_b$ results were then averaged.

Figure 5.5 shows accuracies across all participants for classification time windows up to 1 s. For both approaches, the classifier reached high mean accuracies above 90% for the 1 s time window.

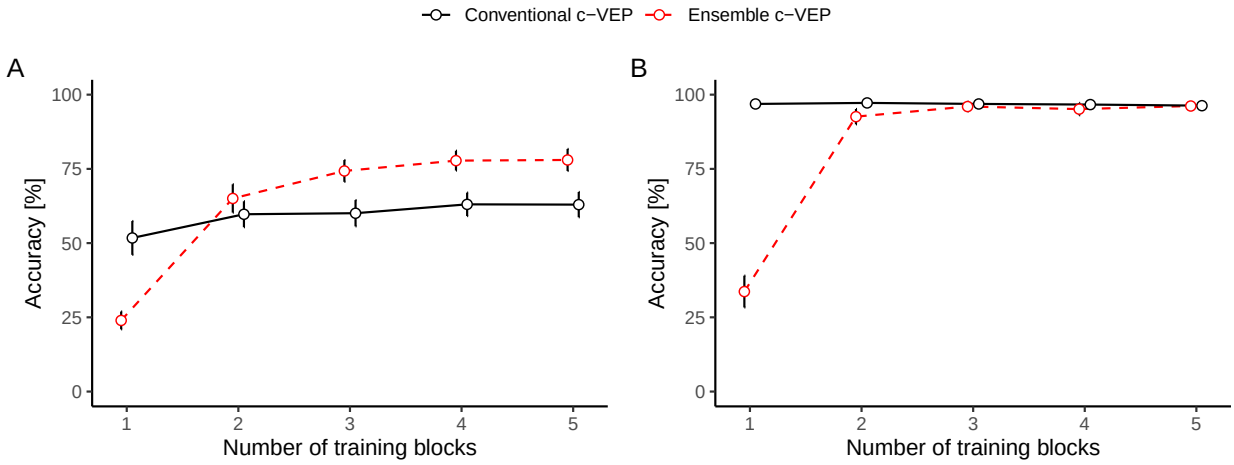



**Figure 5.6:** Mean accuracies across participants with different numbers of training blocks for each method. The red dashed line represents the ensemble-based classification, the solid black line the conventional c-VEP classification. The error bars indicate standard errors of the means. (A) The classification time window was 0.3 s. (B) The classification time window was 1 s.

As expected, the accuracy increased for larger time windows. Nevertheless, it can be seen that for the ensemble-based approach, a time window as low as 0.4 s still yielded accuracies around 90% for most participants, indicating that reliable control can be achieved long before a full stimulation cycle is completed.

To further compare the performance of the two classification methods, the effect of training length (i.e., the number of blocks) on the classification accuracy was investigated. Figure 5.6 shows the classification accuracy for different numbers of training blocks at classification time windows of 0.3 s and 1 s. For the conventional approach, 1 block yields 8 trials. For the ensemble approach, 1 block yields 1 trial per class. The ensemble-based approach generally requires more training data to yield its maximum performance, as each class is recorded separately.

For both methods, the classification accuracy increased when increasing the number of training blocks. For the 0.3 s time window (see Figure 5.6 A), the ensemble method outperformed the standard approach if more than two training blocks were used. For the 1 s time window (see Figure 5.6 B), both methods yielded high classification accuracies. However, the ensemble method required three training blocks to reach mean accuracies above 95%; the conventional method achieved this threshold at all numbers of training blocks.

In contrast to the studies in the previous chapter, we used 16 signal channels instead of only 8. The effect of the number of channels on the classification accuracy and signal-to-noise ratio (SNR) was investigated. For c-VEP signals, the SNR can be estimated by calculating the similarity index (SMI), which describes the power

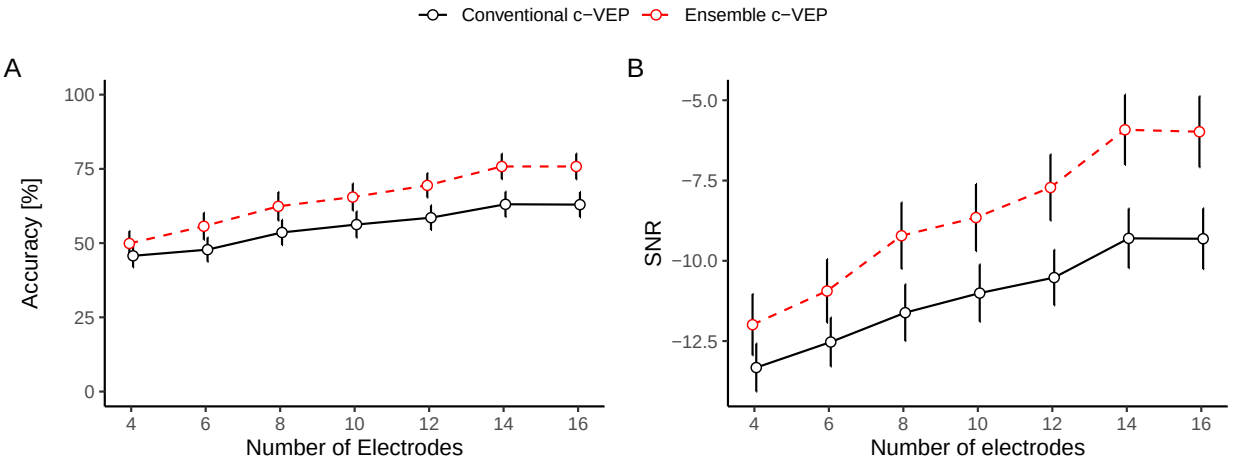



**Figure 5.7:** (A) Mean accuracies and (B) signal-to-noise ratios (SNRs) across participants with different numbers of electrode channels for each method. The red dashed line represents the ensemble-based classification, the solid black line the conventional c-VEP classification. The classification time window was 0.3 s. The error bars indicate standard errors of the means.

ratio between the test trial and the residual (see, e.g., [101]). Let $\bar{\mathbf{X}}$ denote the averaged training signal and let $\mathbf{x}_i$ denote the $i$-th test signal. We then calculate $\mathbf{S}$, the orthogonal projection of $\mathbf{x}_i$ onto $\bar{\mathbf{X}}$, and $\mathbf{N}$, the residual part:

$$\mathbf{S} = \frac{\bar{\mathbf{X}}^T \mathbf{x}_i}{\bar{\mathbf{X}}^T \bar{\mathbf{X}}} \bar{\mathbf{X}} \quad \text{and} \quad \mathbf{N} = \mathbf{x}_i - \bar{\mathbf{X}}. \tag{5.8}$$

The SMI in decibels (dB) is calculated as

$$\text{SMI} = 10 \log_{10}(\sigma^2(\mathbf{S})/\sigma^2(\mathbf{N})), \tag{5.9}$$

where $\sigma^2$ represents the variance. The more similar training and test trials are to each other, the higher the SMI. For this reason, the SMI can be used to approximate SNRs of c-VEP signals [101]. To asses the SNRs, the SMIs were calculated for the spatially filtered training and test signals during off-line cross-validation.

Figure 5.7 shows classification accuracies and SNRs for the ensemble and the conventional classification for different channel montages around the visual cortex. More precisely, the following montages were compared

4: Pz, O1, Oz, O2
6: Pz, O1, Oz, O2, PO3, PO4
8: Pz, O1, Oz, O2, PO3, PO4, O9, O10
10: Pz, O1, Oz, O2, PO3, PO4, O9, O10, P3, P4
12: Pz, O1, Oz, O2, PO3, PO4, O9, O10, P3, P4, POz, PO8
14: Pz, O1, Oz, O2, PO3, PO4, O9, O10, P3, P4, POz, PO8, POO1, POO2
16: Pz, O1, Oz, O2, PO3, PO4, O9, O10, P3, P4, POz, PO8, POO1, POO2, P5, P6.

The channel selection for each montage was based on results from a previous study, where we analyzed all possible combinations for different numbers of electrodes [152]. It can be seen that the classification accuracy and the SNR decrease with the reduction of

**Table 5.2:** On-line performance of the dictionary-driven c-VEP speller. Provided are the information transfer rates (ITRs), accuracies, and the output characters per minute (OCM) for the letter-by-letter spelling task BRAIN and the subject-specific individual sentence task as listed in Table 5.1 (Sent.). Data from Gembler and Volosyak [72].

| Subject | Accuracy [%] | | ITR [bpm] | | OCM [char/min] | |
|---|---|---|---|---|---|---|
| | BRAIN | Sent. | BRAIN | Sent. | BRAIN | Sent. |
| 1 | 100 | 97 | 84.7 | 60.1 | 14.1 | 19.3 |
| 2 | 100 | 96 | 65.9 | 45.5 | 12.2 | 15.1 |
| 3 | 100 | 95 | 79.3 | 57.5 | 13.2 | 20.1 |
| 4 | 100 | 100 | 74.1 | 58.9 | 12.3 | 24.0 |
| 5 | 100 | 100 | 54.9 | 61.6 | 9.1 | 19.7 |
| 6 | 100 | 97 | 71.6 | 58.3 | 11.9 | 18.8 |
| 7 | 100 | 100 | 60.8 | 57.0 | 10.1 | 17.3 |
| 8 | 100 | 100 | 79.8 | 80.0 | 13.3 | 19.0 |
| 9 | 100 | 100 | 97.0 | 48.1 | 16.2 | 15.4 |
| 10 | 100 | 100 | 125.4 | 95.8 | 20.9 | 22.6 |
| 11 | 100 | 97 | 79.0 | 53.7 | 13.2 | 15.0 |
| 12 | 86 | 96 | 43.2 | 49.4 | 7.7 | 19.1 |
| 13 | 100 | 100 | 86.1 | 76.0 | 14.4 | 21.4 |
| 14 | 100 | 86 | 85.1 | 46.8 | 14.2 | 18.5 |
| 15 | 92 | 90 | 50.2 | 49.0 | 8.1 | 13.8 |
| 16 | 100 | 91 | 91.1 | 62.6 | 15.2 | 23.9 |
| 17 | 100 | 100 | 57.2 | 46.1 | 9.5 | 11.4 |
| 18 | 100 | 82 | 77.1 | 34.5 | 12.8 | 16.6 |
| Mean | 98.8 | 95.9 | 75.7 | 57.8 | 12.7 | 18.4 |

the electrodes. Moreover, the accuracies and SNRs were generally larger for the ensemble-based method in comparison to the conventional method. It should be noted that increasing the number of electrodes may decrease user comfort; more time is required for preparation and cleaning.

All participants completed the on-line experiment. The on-line performance for letter-by-letter and sentence spelling tasks were evaluated with the command accuracy, the ITR, and the OCM. The ITR evaluates performance on the level of the target identification. It does not depend on features of the application such as integrated dictionary suggestions; although these features can increase the performance on the application level (e.g., a higher number of characters per minute can be achieved due to word suggestions), they do not affect the ITR (see section 2.7).

Table 5.2 displays the results of the on-line spelling tasks. In terms of detection accuracy, all participants were able to complete the task with average accuracies above 80% for the letter-by-letter and the sentence task. For the letter-by-letter spelling task BRAIN, a mean (SD) accuracy of 98.8 (3.7)% was reached; for the sentence spelling task, a mean accuracy of 95.9 (5.4)% was reached. Sixteen out of the eighteen participants completed the spelling task BRAIN without any errors, reaching an accuracy of 100%. For the sentence spelling tasks, still, eight participants reached 100% classification accuracy.

The mean (SD) ITR for the spelling task BRAIN was 75.7 (19.3) bpm. For the individual sentence spelling task, it was significantly lower, 57.8 (14.4) bpm (paired $t$-test: $t = 4.66$, $p < 0.001$). Across individual participants, the minimal and maximal ITR were 43.2 bpm and 125.4 bpm for the spelling task BRAIN and 34.5 bpm and 95.8 bpm for the sentence spelling task, respectively.

In terms of OCM, significantly better results were achieved when the dictionary integration was used. The average OCM was 12.7 (3.2) char/min for spelling BRAIN and 18.4 (3.5) char/min for the individual sentence task ($t = 6.91$, $p < 0.00001$). Across individual participants, the minimal and maximal OCM were 7.7 char/min and 20.9 char/min for the spelling task BRAIN and 11.4 char/min and 20.4 char/min for the sentence spelling task, respectively.

### Discussion

In this study, we presented a dictionary-driven c-VEP spelling application employing $n$-gram based dictionary suggestions. In contrast to the typically fixed time windows used in most state of the art c-VEP spellers, a dynamic time window mechanism was implemented, which allowed accurate discrimination between intentional and unintentional fixations. If the user did not focus on a particular button or just briefly attended it, e.g., when searching for the desired character, the threshold criterion was not met, and no classification was performed.

Another advantage of the approach is the additional user feedback provided through progress bars. Typically, in c-VEP-based BCIs, feedback is given on trial base only, i.e., after fixed time intervals. This kind of feedback is also referred to as discrete feedback [153]; for example, after each trial of fixed length, the selected letter may be added to an output display. Here, continuous feedback was provided throughout the trial in the form of progress bars, which were updated each calculation interval (here 0.05 s). This real-time information about the classification state is also valuable to customize system parameters during familiarization. Similar methods have been incorporated into asynchronous SSVEP-based BCI systems and lead to increased user-friendliness and system accuracy [42, 154].

The selection options of the GUI changed dynamically. For example, the dictionary suggestions were updated after each selection. Such changing elements of the GUI could be handled easily due to the dynamic time window approach.

For two-step spelling interfaces, like the one presented here, each letter selection comprises of two classifications and two gaze shifting phases. It remains to be tested if the dictionary support is

as beneficial for multi-target systems that require only one step to select a character (see section 5.4 in this chapter).

Another addition to the state-of-the-art is the software-based stimulus onset determination (see section 3.6). The high accuracies achieved in the study demonstrate the reliability of this approach. The same principle can also be adapted to SSVEP systems that employ a hybrid frequency and phase coding stimulus design (this will be addressed in section 5.3).

In addition to the latency of the stimulus presentation, a short time interval elapses between stimulus presentation of the eye and the occurrence of a VEP. Although not applied here, some researchers achieved improvements in BCI performance by excluding samples from the beginning of the data buffer to address the latency of the visual system. For example, Wittevrongel et al. [102] recommended excluding the first 150 ms of the trials from the decoding for the c-VEP paradigm. Similarly, Jia et al. [155] found SSVEP latencies of different stimulus frequencies to be around 130 ms.

As evident from the off-line analysis, the classifier produced accurate labels before a full stimulation cycle was completed. In general, the ensemble-based approach demonstrated superior off-line performance (i.e., higher accuracies for short time windows). This performance boost seems to carry over to on-line performance as well: In our previous study [98], we used the conventional c-VEP classification approach for copy spelling tasks using the same interface; ten participants spelled individual sentences with a mean ITR of 31.1 bpm. Here, the mean ITR was roughly twice as high (i.e., 57.8 bpm).

A downside of the ensemble-based strategy is the prolonged training duration. Performance typically increases when longer training sessions are conducted. Here, we averaged the data over six trials for the ensemble approach. As eight targets were used, the same data yielded 48 trials with the conventional approach.

Recently, Nagel et al. [113] investigated the effect of monitor raster latencies on target detection; the raster latency is dependent on the vertical refresh rate of the screen and causes small additional time lags between vertically separated targets. According to the authors, the correction of these latencies can enhance detection accuracy.

It should further be noted that some c-VEP-BCIs employ additional flickering objects around the selectable targets (principal of equivalent neighbors, see section 3.2) to increase the similarity between trials. This strategy was not applied here, as it involves additional flickering objects which limit GUI design and is visually more demanding.

In general, higher ITRs than reported here can be achieved with the c-VEP paradigm when using multi-target GUIs. For example, the system developed by Spüler et al. [9] achieved 144 bpm and an average of 21.3 error-free letters per minute in on-line spelling tasks; the authors implemented a 32-target c-VEP system with fixed classification time windows of 1.05 s. However, thanks to the dictionary integration, the average number of error-free characters achieved in the presented study (i.e., 18.4 char/min) was quite similar, albeit using only eight targets.

VEP-based BCIs are often compared with eye-tracking interfaces, as both require eye gaze control. A significant advantage of the presented system is that it is not affected by the Midas touch problem. The responsiveness of the presented system was also promising; hence, the c-VEP paradigm could be hybridized, e.g., with eye-tracking technology, as described in our previous publication, where we combined an eye-tracker and an SSVEP application [67].

The dynamic sliding window mechanism, the $n$-gram based dictionary integration, and the implementation of software-based stimulus synchronization used in the reported study add to a growing body of literature on c-VEP-based BCIs.

## 5.2 Effect of Age and Flickering Speed on c-VEP Performance

This section is an amended version of [148]: Gembler et al. (2019), 'A Comparison of cVEP-Based BCI-Performance Between Different Age Groups'.

The strength of the VEP-responses varies across users. Similar to the observations regarding the SSVEP paradigm, system speed is expected to be lower for elderly participants. However, as with SSVEP systems, the majority of studies investigating c-VEP-based BCIs have been conducted with a young subject group.

In the reported study, we investigated the level of fatigue and annoyance for different c-VEP flashing speeds. We tested two equal-sized groups of different age ranges with the eight-target spelling application presented in the previous section.

Next to system speed, overall user-friendliness is another crucial factor in BCI research. As mentioned in section 4.1, the stimulation pattern used to evoke the brain response may be considered as annoying, especially for lower rate flickering. Interestingly, in their SSVEP field study, Allison et al. [82] reported that younger participants tended to be less annoyed by the flickering.

For c-VEP systems, several studies suggest that the flickering rate plays a central role in the perceived level of annoyance and performance [98, 102]. Here, three flickering speeds were investigated;

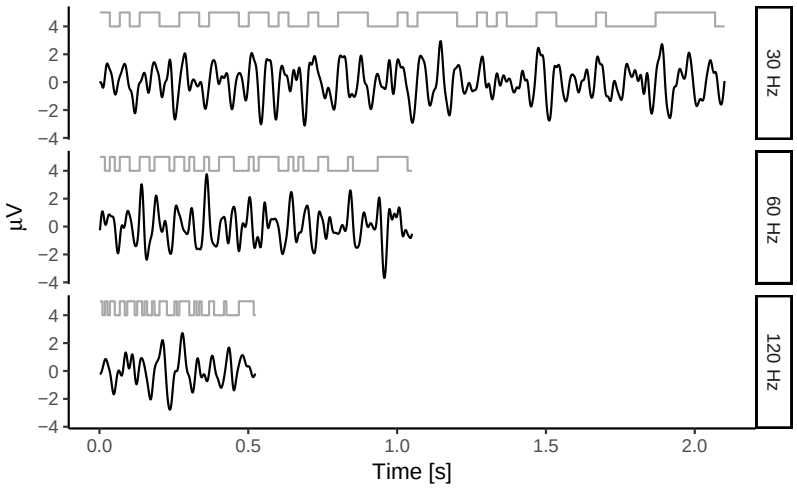



**Figure 5.8:** The 63 bit *m*-sequences (gray) and the reference templates of subject 5 for 30, 60, and 120 Hz setups. For each setup, one stimulus cycle is shown.

the c-VEP stimuli were generated with monitor refresh rates set to 30, 60, and 120 Hz.

## Methods

This section presents details about the subject groups and describes the experimental design. The hardware setup and signal processing methods were identical to the study presented in the previous section. Furthermore, the same spelling interface was used.

For the stimulus design, eight 63 bit *m*-sequences were used (as determined in the previous section). Three setups were tested: The update rate $r$ of the bit pattern was set to 30, 60, and 120 Hz, respectively. The duration of the stimulus cycle, $r/63$, was 2.1, 1.05, and 0.525 s.

### Participants

In total, 26 healthy participants were recruited for this experiment. The participants were divided into two groups based on their age. The group of younger participants (referred to as the young group) had a mean (SD) age of 23.5 (2.6) years, ranging from 20 to 28; five participants were female. The group of elderly participants (referred to as the elderly group) had a mean (SD) age of 72.7 (6.3) years, ranging from 62 to 83; eight participants were female. All subjects had normal or corrected-to-normal vision. Spectacles were worn if needed. The entire session lasted approximately 60 minutes for each subject. All participants received a financial reward for their participation.

### Experimental protocol

The experimental design was the same for both age-groups: The experiment consisted of three sessions were different refresh rates were tested (30, 60, and 120 Hz). Each session consisted of a training phase and a copy spelling phase.

As in the previous section, each training phase was grouped in six training blocks, $n_b = 6$, where $6 \cdot 8 = 48$ trials were collected in total. Each of these trials lasted for 4.2 s, i.e., the stimulation cycle repeated 2, 4, or 8 times depending on the used refresh rate (30, 60, and 120 Hz, respectively).

Before the copy spelling task, a brief familiarization run was conducted were subjects spelled the word BCI, and, if necessary, the classification parameters were adjusted manually to ensure effective control. The copy spelling task (for all sessions) was to spell three German words: BAUM, HAUS, and WELT. Participants needed to correct errors using the UNDO function of the eight-target speller. In each session, after completion of the copy spelling tasks, participants rated the level of the annoyance of the flickering on a 1-5 Likert scale.

### Results

Figure 5.8 shows an example of the reference templates for the 30, 60 and 120 Hz setup for one participant. The reference template at a 30 Hz refresh rate resembles the corresponding code sequence the least as it oscillates much more than the code pattern.

Accuracies and ITRs for all participants are shown in Figure 5.9. The results from the three spelling tasks (BAUM, HAUS, and WELT) were averaged. Differences in ITR between refresh rate setups and between age groups were analyzed using paired $t$-tests and unpaired Welch's $t$-tests, respectively.

The young group achieved mean ITRs of 53.1, 64.0, and 72.9 bpm for the 30, 60, and 120 Hz system, respectively. The differences between refresh rate setups were significant for this age group according to paired $t$-tests: The difference between 30 and 60 Hz was significant ($t = 3.25$, $p = 0.007$), the differences between 60 and 120 Hz ($t = 4.61$, $p < 0.001$) and between 30 and 120 Hz ($t = 6.86$, $p < 0.0001$) were highly significant.

The elderly group achieved mean ITRs of 42.03, 45.32, and 45.74 bpm for 30, 60, and 120 Hz refresh rate, respectively. However, for this age group, the differences in means between the refresh rates were not significant.

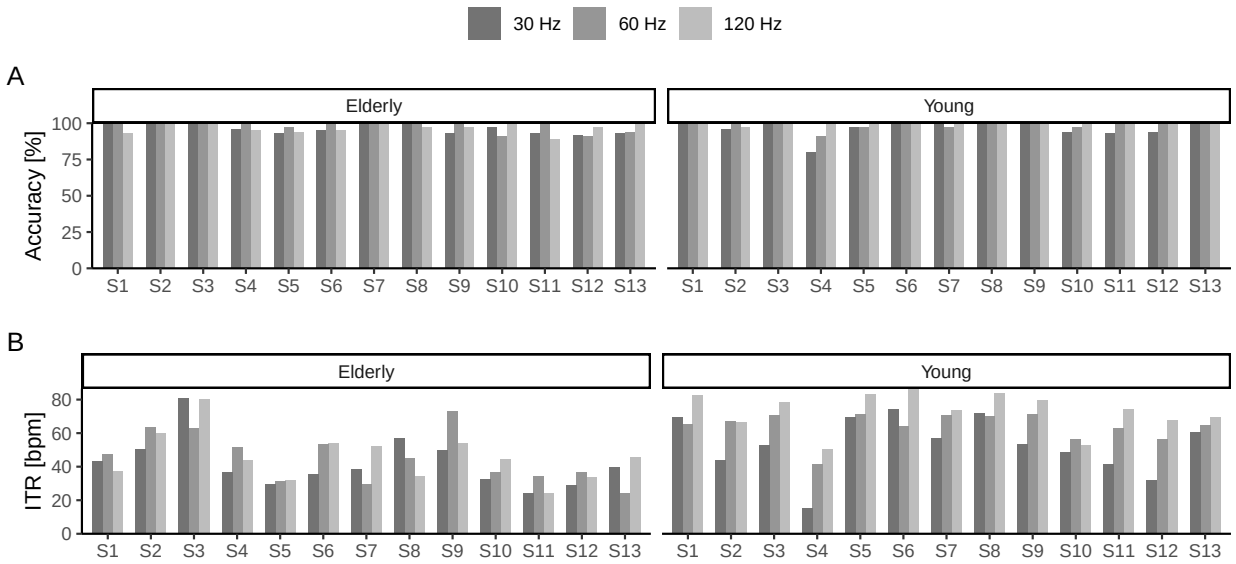



**Figure 5.9:** Individual c-VEP on-line spelling performances of elderly and young participants. The values for the three spelling tasks (BAUM, HAUS, and WELT) were averaged. (A) Accuracies and (B) information transfer rates (ITRs) for 30, 60, and 120 Hz refresh rate setups are shown. Data from Gembler et al. [148].

In regards to the impact of age on BCI performance, Welch's $t$-tests revealed highly significant differences between the mean ITRs of the young and the elderly group for the 120 Hz setup ($t = 5.27, p < 0.0001$) and for the 60 Hz setup ($t = 3.93, p < 0.001$). For the 30 Hz setup, however, no statistical difference between mean ITRs of the young and elderly group was found ($t = 1.75, p = 0.09$). An overview of the age-related performance difference is presented in Figure 5.10.

For all refresh rate settings and age groups, average accuracies above 95% were achieved. Young participants reached mean accuracies of 96.5%, 98.6%, and 99.7%, and elderly participants reached mean accuracies of 96.4%, 97.9%, and 96.6% for the 30, 60, and 120 Hz setup, respectively.

Figure 5.11 shows the results of the user questionnaire. Participants rated their level of annoyance using a five-point Likert scale. For the young group, the median rating was 3 for the 30 Hz (range 1 to 4), 2 for both the 60 Hz (range 1 to 4) and the 120 Hz (range 1 to 4) setup. Similarly, for the elderly group, the median rating was 3 for the 30 Hz (range 1 to 5), and 2 for both the 60 Hz (range 1 to 4) and the 120 Hz (range 1 to 3) setup. Most users did not find the flickering for any refresh rate setup annoying. For the young group, 1 participant out of 13 found the 120 Hz setup annoying, versus 2 for the 60 Hz setup and 4 for the 30 Hz setup. For the elderly group, none of the 13 participants rated the 120 Hz setup annoying or very annoying, versus 1 for the 60 Hz setup and 4 for the 30 Hz setup.

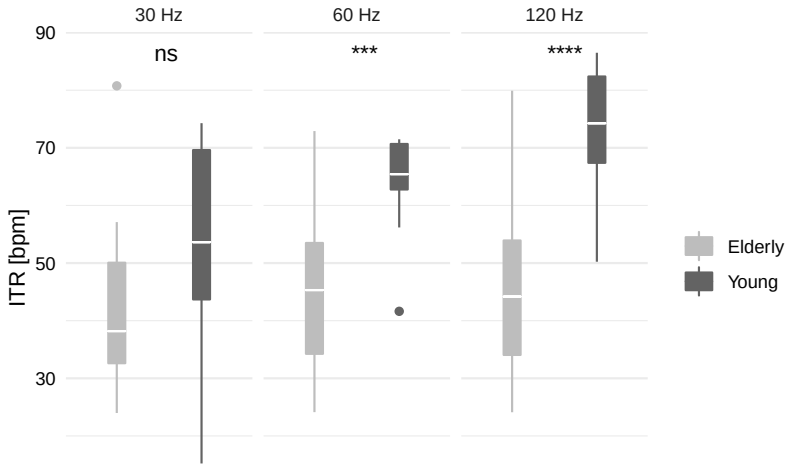



**Figure 5.10:** Comparison of c-VEP online spelling performances of elderly and young participants. The results from young (20 to 28 years) and elderly (62 to 83 years) groups are provided. Displayed are the information transfer rates (ITRs) averaged over the three on-line spelling tasks (BAUM, HAUS, WELT) for three monitor refresh rates that were used to generate the stimuli. The asterisks mark statistical significance (Welch's $t$-test, $***p < 0.001$ and $****p <= 0.0001$).

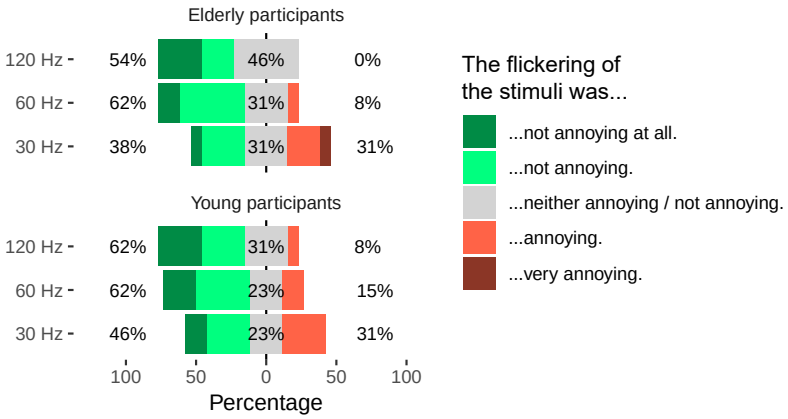



**Figure 5.11:** Subjective level of user-friendliness for young and elderly participants when using a c-VEP-based BCI system. Participants were asked to state the perceived level of annoyance on a 1 to 5 Likert scale.

Overall, answers indicate that the highest flickering speed was perceived as the least annoying. Additional comments from the participants revealed reasons why the 120 Hz setup was perceived as least annoying. One participant found that for the 120 Hz setup, it was easier to concentrate on the letters of the stimuli. Another participant commented that the 120 Hz setup was least fatiguing.

## Discussion

The presented results reveal a significant difference between the performance of young and elderly participants for the c-VEP paradigm. The performance difference in terms of ITR was most visible for the fastest flickering speed, (i.e., during the 120 Hz refresh rate session). The accuracies were high in all sessions for both groups; mean accuracies were above 95% in all sessions for the elderly and the young group. The reason for the lower ITRs of the elderly group is longer average classification times. This is consistent with the results from our findings with the SSVEP paradigm reported in section 4.3. Another explanation for the

higher mean ITR of the younger group could be that the reaction time is shorter for young participants. To reduce this effect, we included a familiarization run, where participants learned the functioning of the GUI. Moreover, the gaze-shifting phase was set to 2 s, which is comparably long but provides enough time to locate the desired target.

Improved BCI performance for faster flickering speeds was also observed by Wittevrongel et al. [102], who reported that a 120 Hz refresh rate enabled higher ITRs in comparison to the standard 60 Hz stimulus presentation. In the presented study, however, a difference between refresh rates was only observed for the young group. The results achieved in this study could indicate that higher refresh rates than tested here could yield even higher ITRs. In preliminary experiments, we also tested with higher flickering speeds (using a refresh rate of 200 Hz). Although the flickering was perceived as less annoying, the 200 Hz yielded, on average, the lowest accuracies. According to the questionnaire results, the flickering speed has an impact on user comfort. The 120 Hz session was rated least annoying in terms of the flickering.

In summary, for the 60 and 120 Hz sessions, a highly significant difference between young and elderly participants was found. Thus the study confirms our results obtained with the SSVEP paradigm suggests that user age needs to be considered when designing c-VEP BCIs. The 120 Hz setup yielded higher ITRs than both the 60 Hz and the 30 Hz setup for the young group. As the 120 Hz setup was considered least annoying by both groups, it seems to be the better option in general.

## 5.3 Automated Calibration and Comparison of c-VEP and SSVEP

This section is an amended version of [149]: Gembler et al. (2019), 'Dynamic Time Window Mechanism for Time Synchronous VEP-Based BCIs—Performance Evaluation with a Dictionary-Supported BCI Speller Employing SSVEP and c-VEP'.

Although SSVEPs and c-VEPs are the most commonly used VEPs in BCI research, a direct side-by-side comparison is so far still missing. It would be interesting to find out, which of the two stimulation paradigms yields the fastest spelling speeds and which paradigm causes the least visual fatigue.

The advantages and disadvantages of the two VEP paradigms have already been touched on briefly in the introduction and the previous chapter. Both paradigms yield high ITRs if individual EEG recordings are used. Although SSVEP-based BCIs can be realized without training sessions, the highest spelling speeds are achieved for both paradigms, if pre-recorded user EEG data are used to classify the attended target [114]. This common factor enables a direct comparison in the sense that the visual flickering

is the main difference between the paradigms; the same number of classes, the same signal classification algorithms based on template matching can be used.

To compare the two paradigms, an experiment with twelve healthy subjects was conducted. After off-line recording phases, word and sentence spelling tasks were performed, once with SSVEP and once with c-VEP. For both paradigms, the eight-target spelling interface was used. In this sense, the sliding window approach (presented in section 5.1) was adopted for SSVEP-BCIs using individual EEG templates; so far, these systems have mainly been implemented as synchronous systems (i.e., the system employs predefined classification time windows). In addition to that, the study explores autonomous calibration of minimal classification time windows and thresholds for both paradigms.

The BCI performance of the two flickering methods was assessed using the standard measures accuracy, OCM, and ITR. The usability was assessed using user questionnaires.

## Methods

In the following, the subject group, the stimulus presentation, the experimental protocol, and the automated parameter setup are described in detail. The hardware setup and signal processing methods were similar to the study presented in section 5.1. Moreover, the dictionary-driven eight-target speller was used.

### Participants

Twelve healthy participants were recruited for this experiment, eight females and four males (average age 23.8 years, SD 2.35, range 21 to 30 years). All participants had normal or corrected to normal vision. The subjects received a financial reward for their participation.

### Stimulus presentation

Two stimulus types, SSVEP- and c-VEP stimuli, were tested consecutively. The target stimuli consisted again of eight boxes (230 $\times$ 230 pixel) arranged as 2 $\times$ 4 stimulus matrix (see section 5.1). Accordingly, the number of stimulus classes, $K$, was set to 8.

For the c-VEP flashing pattern, eight 63 bit $m$-sequences were used (where the initial code $c_1$ was defined as in section 5.1). According to Wei et al. [103], a lag of 4 bit between adjacent stimuli yields better performance for a modulation sequence with a length of

63 bit. Therefore, in this study, the remaining $K - 1$ targets were generated by employing a circular shift of 4 bit ($c_1$ had no shift, $c_2$ was shifted by 4 bit to the left, $c_3$ was shifted by 8 bit, and so on). The duration of one stimulus cycle was 1.05 s. As the sampling rate was set to 600 Hz, the number of samples per cycle, $n_c$, was 630 samples.

For the SSVEP flashing pattern, hybrid frequency and phase coding was used (see section 3.1). A specific frequency $f$ and phase $\Phi$ were assigned to each target. The flickering was realized by sinusoidally modulating their transparencies in accordance with the frequency and phase combination, as described in section 3.1. Frequencies $f_i = f_0 + (i-1)\Delta_f$ and phases $\Phi_i = \Phi_0 + (i-1)\Delta_\Phi$, $i = 1, \ldots, K$, with $f_0 = 8\,\text{Hz}$, $\Delta_f = 1\,\text{Hz}$, $\Phi_0 = 0$ and $\Delta_\Phi = 0.35\pi$, where assigned column-wise to the stimulus matrix. The frequency range from 8 to 15 Hz was chosen, as it avoids mutual influences between fundamental and harmonic frequencies (as discussed in section 4.1). Furthermore, due to the 1 Hz difference between stimuli, the stimulus repetition period is 1 s.

**Experimental protocol**

Participants sat on a chair facing the LCD screen (at a distance of approximately 60 cm). After they were prepared for the EEG recording, they went through two sessions (c-VEP and SSVEP). Each session consisted of a training phase (for template recording and automated parameter setup), an on-line copy spelling phase, and a brief questionnaire. The experiment took approximately one hour for each participant. The order of the paradigms was altered for every other participant. Hence half of the participants started the experiment with the SSVEP paradigm, the other half with the c-VEP paradigm.

**Training phase** In the training phase, the user fixed his or her gaze at each of the eight stimuli several times. For each trial, the code pattern repeated for three cycles, i.e., the stimuli flickered for $3 \cdot 1\,\text{s} = 3\,\text{s}$ for the SSVEP paradigm and for $3 \cdot 1.05\,\text{s} = 3.15\,\text{s}$ for the c-VEP paradigm. A green frame indicated which box the user needed to fixate. The recording was grouped in six blocks ($n_b = 6$) of eight trials, resulting in $6 \cdot 8 = 48$ trials in total. To avoid user fatigue, the experimenters allowed the subjects to take breaks after each block of eight trials (the recording automatically paused). To start another recording block, the subjects needed to press the space bar.

**Table 5.3:** Individual sentence tasks of the SSVEP and c-VEP on-line experiments. The mean sentence length was 26.7 and 26.8 characters for the c-VEP and SSVEP sentences, respectively.

| # | Sentence c-VEP | Sentence SSVEP |
|---|---|---|
| 1 | THE TRAIN WAS OVERCROWDED | HE IS AFRAID OF HORSES |
| 2 | THAT WAS AN INTERESTING LECTURE | I WANT TO BECOME A BUSDRIVER |
| 3 | THE BOOK IS WAY TOO BORING | I WOULD LIKE TO PLAY THE CELLO |
| 4 | I DO NOT LIKE TO EAT FISH | I DO NOT SPEAK FINNISH |
| 5 | DID YOU EVER DRIVE A SKATEBOARD | ALL OF THE PHOTOS WERE BLURRED |
| 6 | THE RECORDING IS REALLY BAD | I COULD EAT PIZZA EVERYDAY |
| 7 | I DO NOT LIKE THIS MUSIC AT ALL | I USUALLY FALL ASLEEP IN THE CINEMA |
| 8 | DOGS ARE NOT ALLOWED | THAT WAS A NICE MOVIE |
| 9 | I WILL GO SWIMMING TOMORROW | THEY OWN A BLACK CAT |
| 10 | I NEED TO BUY A NEW TOOTHBRUSH | MY BIKE HAS NOT BEEN STOLEN |
| 11 | THE SHOP WAS CLOSED ALREADY | I WANT TO LISTEN TO THE RADIO NOW |
| 12 | THE DOG BARKED LOUDLY | DID YOU GO TO SCHOOL TODAY |

**Copy spelling phase** Prior to the copy spelling task, a brief familiarization run was performed, where participants spelled the word KLEVE, and a word of free choice (e.g., the own first name). During this familiarization run, in some cases, the automatically determined classification thresholds were lowered manually to increase the responsiveness of the application. In the copy spelling phase, participants were first asked to spell the words BCI and BRAIN, and then a longer English sentence. For each participant and paradigm, different sentences were used (see Table 5.3). Occurring errors were corrected using the UNDO function of the interface.

**Questionnaires** Before the training phase, participants filled in a brief questionnaire, answering questions regarding gender and age. Additionally, after each session, participants gave their subjective impressions of the BCI answering questions regarding fatigue and annoyance. The questions and the collected answers from these questionnaires are provided in the results section.

### Classification

Spatial filtering on the basis of the training data and ensemble-based target identification as described in section 5.1 were used for classification; ensemble correlations, $\lambda_k$, were determined by stacking all target-specific spatially filtered data and template vectors.

Additionally, the difference between target- and non-target correlations can be enhanced further by applying a filter bank method, which decomposes VEP-data in sub-band components as described

in [68]. The lower and upper cut-off frequencies for the $m$-th sub-band were selected as $m \cdot 8$ and 60 Hz. To this end, an 8-th order Butterworth filter was employed. Forward and reverse filtering was used to cancel the phase response [156].

The ensemble approach was then applied to each sub-band component individually, yielding a set of correlations $\lambda_k^{(1)}, \lambda_k^{(2)}, \ldots, \lambda_k^{(M)}$, $k = 1, \ldots, K$, where $M$ denotes the number of considered sub-bands. Then, the output command candidate was determined using weighted linear combinations of the correlations,

$$C = \underset{k=1,\ldots,K}{\arg\max}\, \tilde{\lambda}_k, \text{ where } \tilde{\lambda}_k = \sum_{m=1}^{M} a_m \lambda_k^{(m)}. \tag{5.10}$$

Mirroring the decrease in amplitude in the higher bands, the weights $a_m$ in (5.10) were set to

$$a_m = \frac{a'_m}{\sum_{k=1}^{M} a'}, \text{ with } a'_m = m^{-1.25} + 0.25, \tag{5.11}$$

yielding decreasing weights for the higher bands. The optimal choice of these weights needs to be investigated further (see, e.g. [68]). For the c-VEP paradigm, the number of sub-bands, $M$, was set to 1 (the standard method), whereas for the SSVEP paradigm, $M$ was set to 5.

After an output command was produced, a gaze shifting phase of 1 s followed. In this gaze shifting period, the amplifier data blocks were ignored, and the stimuli did not flicker, allowing the user to shift his or her gaze to the next target.

**Automatic parameter calibration** For the on-line copy spelling phase, the threshold-based sliding window mechanism was used, where the decision certainty, $\Delta_C$, was defined as the distance between the highest and second-highest correlation. The output was performed if $n_y \geq n_{y_{\min}}$ and $\Delta_C \geq \beta$, where $n_y$ denotes the length of the dynamic time window in samples and $n_{y_{\min}}$ denotes the minimal time window.

Here, the values for the classification threshold $\beta$ and for the minimum time window $n_{y_{\min}}$ were determined automatically for each participant on the basis of the training data via a cross-validation (see section 2.7).

Using stratified 6-fold cross-validation on the training data, an average ITR was calculated after each calculation interval, i.e., for classification windows of $n_y = 30, 60, \ldots, 3n_c$ samples. The value of $n_y$ that maximized the ITR was selected as the minimum time window $n_{y_{\min}}$. The classification threshold $\beta$ was selected as the

| Subject | Time window [s] | | Off-line ITR [bpm] | |
|---|---|---|---|---|
| # | c-VEP | SSVEP | c-VEP | SSVEP |
| 1 | 0.50 | 1.35 | 98.0 | 76.6 |
| 2 | 0.35 | 0.30 | 126.7 | 131.6 |
| 3 | 0.40 | 0.45 | 122.2 | 124.1 |
| 4 | 0.60 | 0.90 | 82.4 | 94.7 |
| 5 | 0.50 | 0.45 | 108.1 | 118.0 |
| 6 | 0.55 | 0.80 | 104.6 | 95.0 |
| 7 | 1.00 | 0.45 | 76.6 | 71.1 |
| 8 | 0.45 | 0.65 | 111.8 | 79.9 |
| 9 | 0.90 | 0.95 | 63.5 | 66.2 |
| 10 | 0.80 | 0.45 | 85.1 | 72.5 |
| 11 | 0.35 | 0.25 | 120.1 | 117.6 |
| 12 | 0.45 | 0.45 | 111.8 | 95.2 |
| Mean | 0.57 | 0.62 | 109.0 | 95.2 |

**Table 5.4:** Off-line comparison c-VEP and SSVEP. Provided are the optimal time window determined in the training (via cross-validation) and the corresponding training ITR for each participant. Data from Gembler et al. [149].

minimal decision certainty, $\Delta_C$, at that time window. An example of the parameter setup procedure is depicted in Figure 5.12.

As stated in the experimental protocol, the suggested thresholds were sometimes lowered manually in the on-line spelling tasks. An explanation for lower certainty in on-line tasks is that in comparison to the cue guided training, it was more likely that participants were not yet gazing at the target when the flickering started.

## Results

**Off-line performance** Optimal time windows and ITRs were calculated via cross-validation. As expected, the highest ITRs were achieved with different time windows for each user. The time window yielding maximal ITR, which was used as the minimal time window in the on-line experiment, and the corresponding maximum ITR are listed in Table 5.4. On average, participants reached a theoretical maximal ITR of 100.9 bpm and 95.2 bpm with the optimal time window for the c-VEP and SSVEP paradigm. However, the difference between the paradigms was not statistically significant according to a paired $t$-test ($t = 1.49$, $p = 0.17$). Moreover, Figure 5.13 shows off-line ITRs and accuracies across all participants for a time window length up to 1 s with an interval of 0.1 s.

**On-line spelling performance** The spelling results across all participants for each paradigm are summarized in Table 5.5. For the single word spelling tasks (BCI and BRAIN), the average ITR was 92.7 and 75.1 bpm for the c-VEP and SSVEP system; the difference between the paradigms was significant ($t = 2.50$,

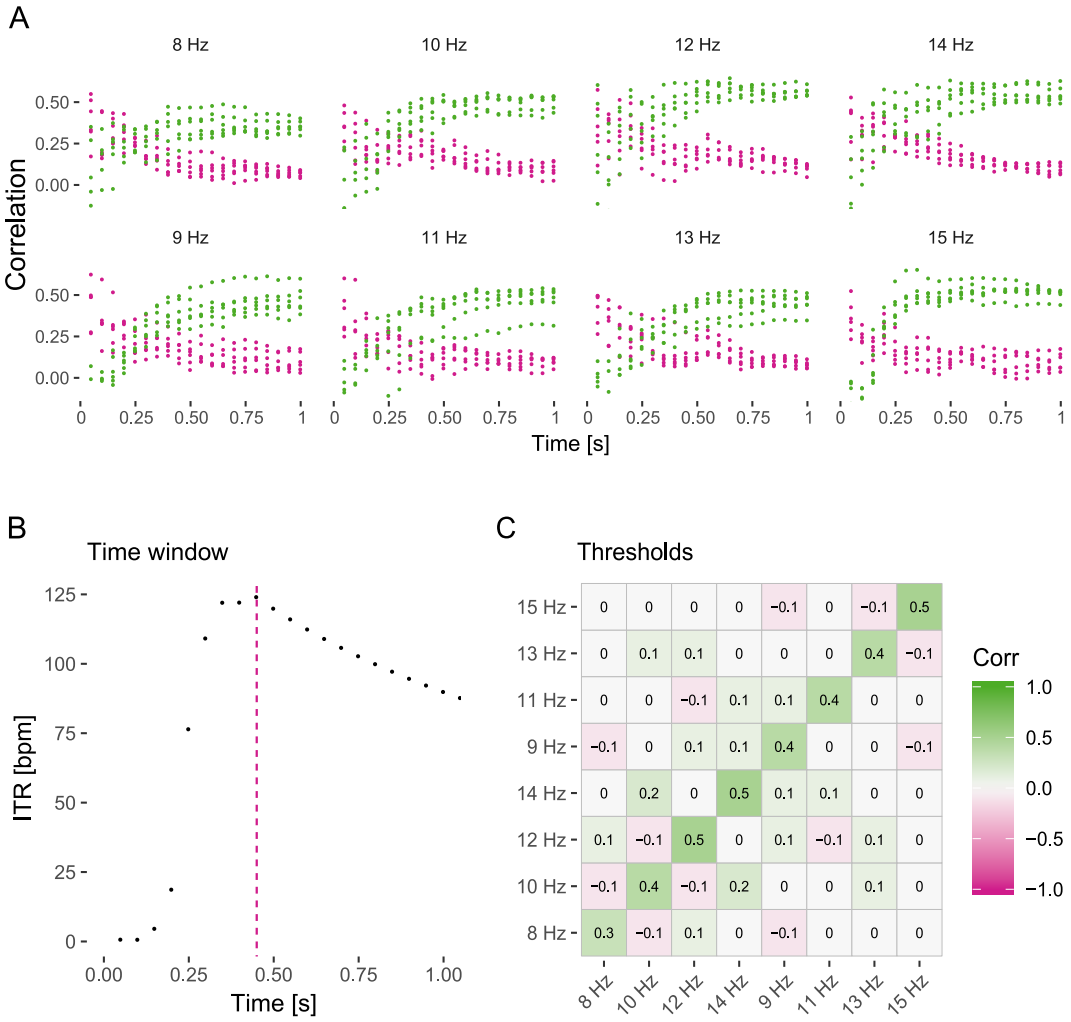



**Figure 5.12:** Example of the automated parameter setup. The figure displays the results from the off-line cross-validation of the SSVEP training data for one participant. In the training session, each of the 8 targets was attended 6 times for 3 s. (A) The correlation values of the target stimulus (green) and the maximum correlation of the non-target stimuli (red) for each stimulus class for time windows up to 1 s. (B) ITR averaged over the 6 training blocks. The dashed line indicates the time window yielding the highest ITR, here 0.45 s. (C) Correlogram of the training data. Depicted are the correlations of each target for the determined time window, averaged over the trials. The classification threshold was determined as the minimal difference between target and non-target stimuli correlation; Here, the distance was minimal for the 10 and 14 Hz pair, yielding a difference of 0.4-0.2=0.2. Adapted from Gembler et al. [149].

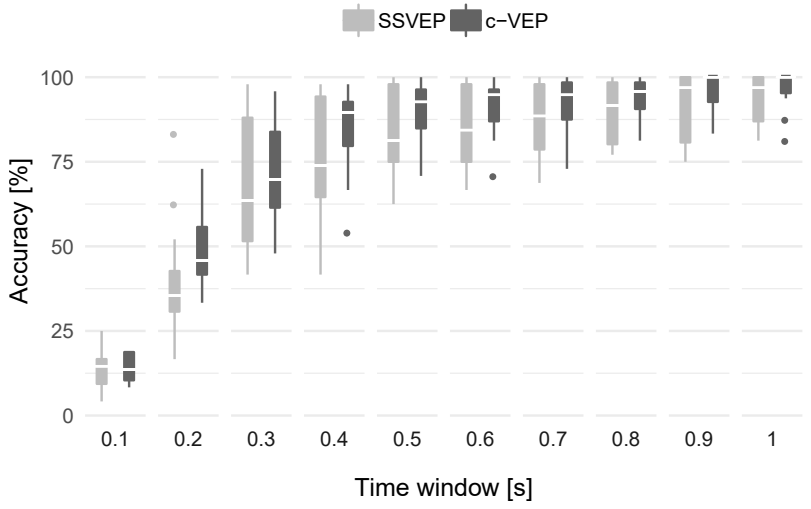



**Figure 5.13:** Classification accuracies achieved with the c-VEP and the SSVEP paradigm. In the box plots, outliers (data points outside 1.5 times the interquartile range) are located outside the "whiskers".

**Table 5.5:** On-line performance comparison c-VEP and SSVEP. Provided are the results (a) for the subject specific individual sentence task as listed in Table 5.3 and (b) for the letter-by-letter spelling tasks BRAIN and BCI. Listed are the accuracies, information transfer rates (ITRs), and output characters per minute (OCM). Data from Gembler et al. [149].

| | **Letter-by-Letter** | | | | | | **Sentence** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | c-VEP | SSVEP | c-VEP | SSVEP | c-VEP | SSVEP | c-VEP | SSVEP | c-VEP | SSVEP | c-VEP | SSVEP |
| Subject | ACC [%] | | ITR [bpm] | | OCM [char/min] | | ACC [%] | | ITR [bpm] | | OCM [char/min] | |
| 1 | 100 | 100 | 100.6 | 32.8 | 16.8 | 5.5 | 97 | 100 | 83.9 | 42.2 | 20.8 | 11.9 |
| 2 | 100 | 87 | 76.5 | 70.7 | 12.7 | 12.3 | 100 | 100 | 36.2 | 62.4 | 12.1 | 16.2 |
| 3 | 100 | 100 | 107.8 | 98.8 | 18.0 | 16.5 | 100 | 100 | 62.1 | 78.6 | 17.9 | 32.8 |
| 4 | 100 | 100 | 85.9 | 45.5 | 14.3 | 7.6 | 96 | 100 | 58.6 | 53.0 | 24.1 | 18.5 |
| 5 | 96 | 100 | 101.7 | 92.5 | 16.8 | 15.4 | 92 | 95 | 49.1 | 51.2 | 16.4 | 14.2 |
| 6 | 100 | 100 | 107.6 | 91.2 | 17.9 | 15.2 | 94 | 93 | 59.9 | 41.1 | 19.0 | 10.2 |
| 7 | 100 | 100 | 69.8 | 73.2 | 11.6 | 12.2 | 92 | 95 | 43.8 | 47.7 | 15.9 | 17.1 |
| 8 | 100 | 96 | 111.0 | 76.7 | 18.5 | 12.6 | 100 | 100 | 89.6 | 77.3 | 37.3 | 28.5 |
| 9 | 94 | 96 | 66.3 | 32.7 | 10.9 | 5.6 | 84 | 91 | 44.0 | 30.4 | 16.7 | 8.0 |
| 10 | 96 | 93 | 84.2 | 88.4 | 13.9 | 15.7 | 88 | 93 | 46.1 | 63.4 | 13.2 | 26.0 |
| 11 | 100 | 100 | 111.9 | 135.3 | 18.7 | 22.5 | 100 | 95 | 64.8 | 96.4 | 20.8 | 22.5 |
| 12 | 96 | 100 | 88.5 | 62.9 | 15.1 | 10.5 | 85 | 94 | 47.3 | 43.4 | 12.3 | 14.1 |
| Mean | 98.5 | 97.6 | 92.7 | 75.1 | 15.4 | 12.6 | 94.0 | 96.3 | 57.1 | 57.3 | 18.9 | 18.3 |

$p = 0.029$). It should be noted that these values were lower than the off-line ITRs as on-line outputs were only produced if the threshold criterion was met.

For the sentence spelling tasks, no statistically significant difference between the paradigms was found; ITRs of 57.1 and 57.3 bpm were achieved for the c-VEP and SSVEP paradigm.

Thanks to the integrated $n$-gram prediction model, the average OCM achieved for the sentence spelling tasks was higher in comparison to the letter-by-letter spelling tasks for both paradigms. For the single word spelling tasks, 15.4 and 12.6 char/min were achieved; for the sentence spelling tasks, 18.9 and 18.3 char/min were achieved with the c-VEP and SSVEP paradigm.

**Questionnaire results** The results from the questionnaires are depicted in Figure 5.14. Most users did not find the flickering annoying or fatiguing. Overall, answers regarding the user-friendliness were slightly more favorable for the SSVEP paradigm. Five out of twelve participants stated that they found the flickering of the c-VEP system annoying. Only one participant found the SSVEP flickering annoying. In respect to the subjective level of fatigue, the SSVEP paradigm yielded better results as well. Four participants found the c-VEP flickering fatiguing, but only two participants stated that the SSVEP flickering caused fatigue.

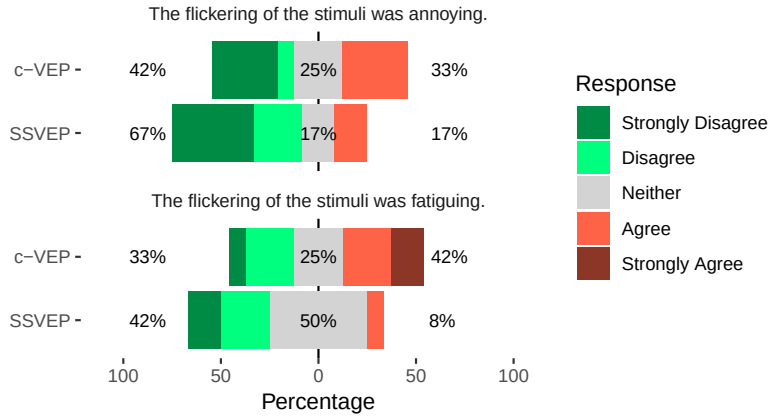



**Figure 5.14:** Subjective level of user-friendliness for SSVEP and c-VEP paradigm. Responses were given on a 1-5 Likert scale, 1 indicating strong disagreement and 5 indicating strong agreement.

## Discussion

The study provides a direct comparison between c-VEP and SSVEP stimulation, both in terms of performance and user-friendliness. The *n*-gram based word suggestion module and the sliding window mechanism were tested with both paradigms.

The comparison of the two stimulation approaches indicates that c-VEP slightly outperforms SSVEP in terms of ITR (see Figure 5.13), but SSVEP is preferred by most users in terms of user-friendliness (see Figure 5.14).

For the SSVEP paradigm, the flickering was realized by sinusoidally modulating the transparencies. This approach allowed a slightly more subtle visual stimulation in comparison to the c-VEP flickering, which switched from full illumination to no illumination of the target in correspondence to the code patterns. Indeed, a slight difference regarding the subjective level of annoyance and fatigue is evident from the questionnaires (see Figure 5.14). In general, most participants seemed to favor the SSVEP paradigm. An even more subtle stimulation could be achieved with motion-based targets using the SSMVEP paradigm [50]. Another approach to reduce the flickering sensation was tested by Chien et al. [157], who employed a composition of red/green/blue 32 Hz/40 Hz flashing lights Chien et al. [157].

The variability across subjects seems to be slightly higher for the SSVEP paradigm than for the c-VEP paradigm (see Figure 5.13). An explanation for this could be that the utilized SSVEP stimulation frequencies interfere more with the natural brain activity (see section 4.1).

A training session was conducted for both paradigms for template recording and parameter optimization. It should be noted that in general, c-VEP-BCIs require a training stage to obtain templates.

SSVEP-BCIs, on the other hand, can be realized without training (i.e., training free) using sine and cosine templates.

One of the critical parameters for BCI performance is the time window used for the classification of the signals. User variability justifies the user-dependent selection of a minimum classification time interval [86]. SSVEP-BCIs have been used with time windows as low as 0.3 s [52]. In some studies, larger classification windows were incorporated to improve the robustness of the system. For example, to outbalance the lower signal-to-noise-ratio with dry electrodes, Spüler [25] used larger classification windows in a c-VEP-BCI by averaging over multiple trials. Similarly, in the previously described SSVEP studies, we incorporated large classification windows to handle age-related inter-subject variability in users [62, 63].

Here, the minimal time window was set user-specifically based on the ITR using the recorded training data. A similar approach was used in our SSVEP wizard (see section 4.4, where the time window was determined with off-line accuracies).

On average, the optimal classification time window in terms of off-line ITR was 0.57 s (ranging from 0.35 s to 1 s) for the c-VEP paradigm and 0.62 s (ranging from 0.25 s to 1.35 s) for the SSVEP paradigm. Hence, despite the comparably low number of targets, high ITRs were achieved. This can also be attributed to the use of ensemble methods, which can significantly increase system speed [52]. Additionally, for the SSVEP paradigm, a filter bank approach, as proposed by Chen et al. [68], was used to enhance target discrimination. Similar methods could also enhance the classification accuracy for the c-VEP approach. In a follow-up study, which is presented in section 5.4, we applied filter bank methods to a multi-target c-VEP system.

Regarding the copy spelling phase, the OCM was, on average, higher for the sentence spelling tasks, in comparison to the word spelling tasks where no dictionary suggestions were used (see Table 5.5). These results confirm the robustness of the time window mechanism and the effectiveness of the implemented word suggestion module. It should be noted that participants did not always use the dictionary whenever they had the chance to do so (suggestions were overseen). Therefore, in some cases, the single word OCM was higher in comparison to the sentence task. Improvements could be made regarding the arrangement of the GUI targets to make the suggestions more prominent. For some participants, a longer gaze shifting phase when suggestions are presented could also be helpful. It should further be noted that participants used the system for the first time. More experience with the GUI could improve the OCM as well.

Overall, there was surprisingly little performance difference between the two stimulation modalities. The c-VEP stimulation patterns yielded slightly higher off-line ITRs and significantly higher ITRs in word copy spelling tasks. In on-line sentence spelling, the speed difference becomes negligible, as usually larger search phases are required to locate the next letter or word.

The results suggest that the stimulation pattern (SSVEP or c-VEP) could be selected based on user preference. In terms of speed, the optimal paradigm could be determined individually for each user in a short training session. However, as the evaluation of the questionnaires suggests, the perceived level of user-friendliness should also be taken into account, as it might be more relevant for end-users than pure system speed.

## 5.4 Investigating Multi-target c-VEP-based BCI Performance

This section is an amended version of [150]: Gembler et al. (2019), 'A Multi-Target c-VEP-Based BCI Speller Utilizing n-Gram Word Prediction and Filter Bank Classification'.

The eight-target system presented in section 5.1 yielded high accuracies in on-line tests. However, in terms of ITR, higher values can be achieved with the c-VEP paradigm by increasing the number of targets. In this section, we applied the methods discussed at the beginning of the chapter to a multi-target interface. Users might need less time to get familiar with the letter arrangement if a typical keyboard layout such as the QWERTZ layout is used; for this, at least 26 classes for letters and additional classes for correction options need to be incorporated.

Several research groups tested 32-target systems resulting in very high ITRs [9, 46]. As for the SSVEP paradigm, a reduction of classification accuracy is expected when using multi-target systems. For example, Bin et al. [46] tested a 16-target and a 32-target system and reported a drop in accuracy from 92% to 85% when using the higher number of targets.

Despite that, novel classification approaches used in SSVEP-BCIs might also boost multi-target c-VEP classification accuracy. Recently, Chen et al. [68] proposed a classification method based on a filter bank design. This approach applies multiple band-pass filters to the recorded EEG to separate the signal into multiple components. Each of these components consists of frequency sub-bands of the originally recorded signal.

Chen et al. [68] used a filter bank method for the SSVEP paradigm. The recorded SSVEP response was segmented with respect to the frequency range of the application; several different designs for the filter bank sub-bands were investigated in their study: equally spaced bandwidths, harmonic frequency bands, and overlapping

sub-bands covering several harmonic frequency bands. From these approaches, the latter resulted in the highest accuracy. Using this method in an on-line experiment, the authors reported an average ITR of 151.2 bpm, which was a significant boost in performance in comparison to standard methods.

Here, we adopted the concept for a c-VEP application. As discussed in section 4.4, closing the eyes during the use of the BCI can lead to false classifications due to alpha activity. To make the system more robust against interferences with naturally occurring brain activity, we decomposed the original signal into alpha-band (8-12 Hz), beta-band (approx. 12-30 Hz), and gamma-band (>30 Hz) related activity.

This approach was tested with a 32-target spelling application. In addition to that, we integrated the sliding window mechanism and the $n$-gram based dictionary module. An on-line experiment with 18 healthy participants was conducted.

## Methods

This section describes the methods and materials used in this experiment. The sliding window mechanism and $n$-gram based dictionary integration, as introduced in section 5.1, were implemented. Also, in terms of hardware, the setup presented in section 5.1 was used.

### Participants

Eighteen healthy participants (eight female, ten male) with a mean (SD) age of 23.6 (4.0) years, range 19 to 31 years, participated in the study. All participants had normal or corrected-to-normal vision. The participants received a small financial reward for participation in this study.

### Stimulus design

The spelling application used 32 boxes (230 × 230 pixel) as stimuli, corresponding to $K = 32$ classes (see Figure 5.15). Each box alternated between 'black' (represented by '0') and 'white' (represented by '1') according to the 63 bit $m$-sequences. The initial code, $c_1$, was defined as in section 5.1. The remaining 31 codes $c_i$, $i = 2, \ldots, K$ were generated by circularly shifting $c_1$ by $(i-1) \cdot 2$ bit to the left. Using this stimulus design, a QWERTZ layout presenting 32 targets (26 letters, 1 underscore, 1 full stop, 3 dictionary suggestions, 1 correction option, as shown in Figure 5.15) was implemented.

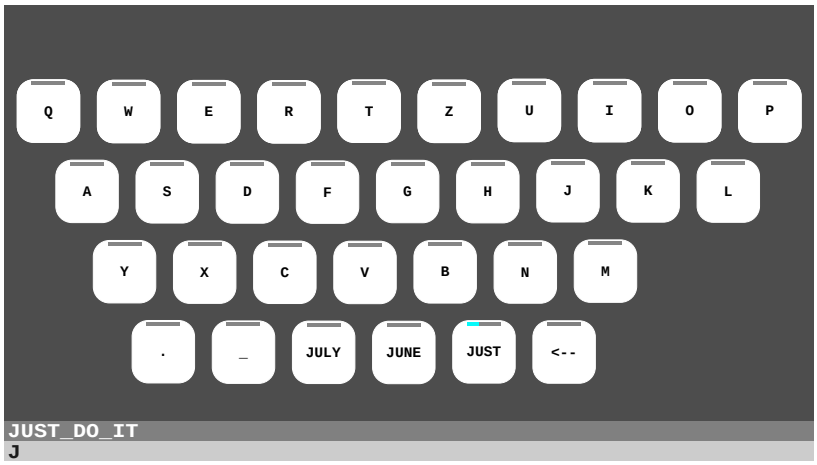



**Figure 5.15:** Interface of the QWERTZ spelling application. The word suggestions based on the $n$-gram prediction model were provided in the lowest row.

### Experimental procedure

Participants went through a recording phase, where the data for the templates and the generation of spatial filters were collected. This session was grouped in three blocks, $n_b = 3$; during each block, the user gazed at each target for 2.1 s (two cycles, where the refresh rate was 60 Hz). In total, $n_b \cdot K = 96$ trials were collected.

The target the user needed to gaze at was highlighted by a green frame. The boxes were highlighted in sequence (upper left to lower right). A gaze shifting phase of 1 s between trials was implemented. After each block, the user could rest until he or she decided to initiate the flickering and recording of the next block by pressing the space bar.

For the generation of CCA-based spatial filters, the standard approach was applied (see also section 5.1). All training trials were shifted to be aligned in phase with the first trial. The shifted trials $\mathbf{Z}_i$, $i = 1, \ldots, n_b K$ were averaged yielding a template $\mathbf{Z}$. Following the standard approach described in section 5.1, a filter vector $\mathbf{w}$ was determined and for each class, templates, $\mathbf{X}_i$, $i = 1, \ldots, K$ were generated by circular shifting $\mathbf{Z}$ in accordance with the bit-shift of the corresponding code $c_i$. Note that in this experiment, templates and weights were determined for $M = 3$ different filter banks using the procedure above. In this respect, $M$ different band-pass filters (described in the following section) were applied to the recorded trials resulting in weights $\mathbf{w}^{(m)}$ and templates $\mathbf{X}_i^{(m)}$, $i = 1, \ldots, K$ for $m = 1, \ldots, M$.

In the on-line session, a brief familiarization run was conducted, where participants learned the functioning of the system. After that, two copy spelling tasks were performed: First, the word BRAIN was spelled (letter-by-letter task), and second, an individual sentence (see Table 5.6) was spelled. Occurring misclassification needed to be corrected by gazing at the box representing the UNDO function.

**Table 5.6:** Individual sentence tasks of the on-line experiment.

| # | Sentence | # | Sentence |
|---|---|---|---|
| 1 | WHAT TIME IS IT | 10 | THE YEAR PASSED BY SO QUICKLY |
| 2 | THE BEER IS IN THE FRIDGE | 11 | I WANT TO LISTEN TO THE RADIO NOW |
| 3 | FILL IN SOME MORE WINE | 12 | DO YOU HAVE A FREE ROOM |
| 4 | CAN I HAVE YOUR NUMBER | 13 | HE OWNS A YELLOW BICYCLE |
| 5 | MY BIKE HAS NOT BEEN STOLEN | 14 | THE SHOP WAS CLOSED ALREADY |
| 6 | DO YOU THINK THAT IS ENOUGH | 15 | MY FAVOURITE COLOUR IS BLUE |
| 7 | WHAT KIND OF MUSIC DO YOU LIKE | 16 | THEY OWN A BLACK CAT |
| 8 | IS THIS SEAT TAKEN | 17 | I WILL GO SWIMMING TOMORROW |
| 9 | JUST DO IT | 18 | WOULD YOU LIKE TO HAVE ICE CREAM |

### Filter bank classification method

Three filter banks were designed using 8-th order Butterworth band-pass filters. The upper and lower cut-off frequencies were set up as follows:

- the first sub-band covered the alpha, beta and gamma bands (a band-pass filter between 8 and 60 Hz was applied);
- the second sub-band covered the beta and gamma band (a band-pass filter between 12 and 60 Hz was applied); and
- the third sub-band covered the gamma band (a band-pass filter between 30 and 60 Hz was applied).

For classification, correlations between the spatially filtered reference signals and the spatially filtered EEG data buffer were calculated for each sub-band ($m = 1, \ldots, M$) independently, yielding a set of correlations

$$\tilde{\lambda}_k^{(m)} = \rho \left( {\mathbf{Y}^{(m)}}^T \mathbf{w}^{(m)}, {\mathbf{R}_k^{(m)}}^T \mathbf{w}^{(m)} \right), \quad k = 1, \ldots, K, \tag{5.12}$$

which were than averaged over the number of filter banks (here $M = 3$),

$$\lambda_k = \frac{1}{M} \sum_{m=1}^{M} \tilde{\lambda}_k^{(m)}, \quad k = 1, \ldots, K. \tag{5.13}$$

The class label $C$ was then determined as $C = \arg\max_{k=1,\ldots,K} \lambda_k$. For the on-line classification, a sliding window mechanism, as described in section 5.1, was implemented.

### Results

The off-line training data were evaluated using 3-fold stratified cross-validation. More precisely, each block (32 trials) was left out once for testing, while the remaining two blocks (64 trials) were

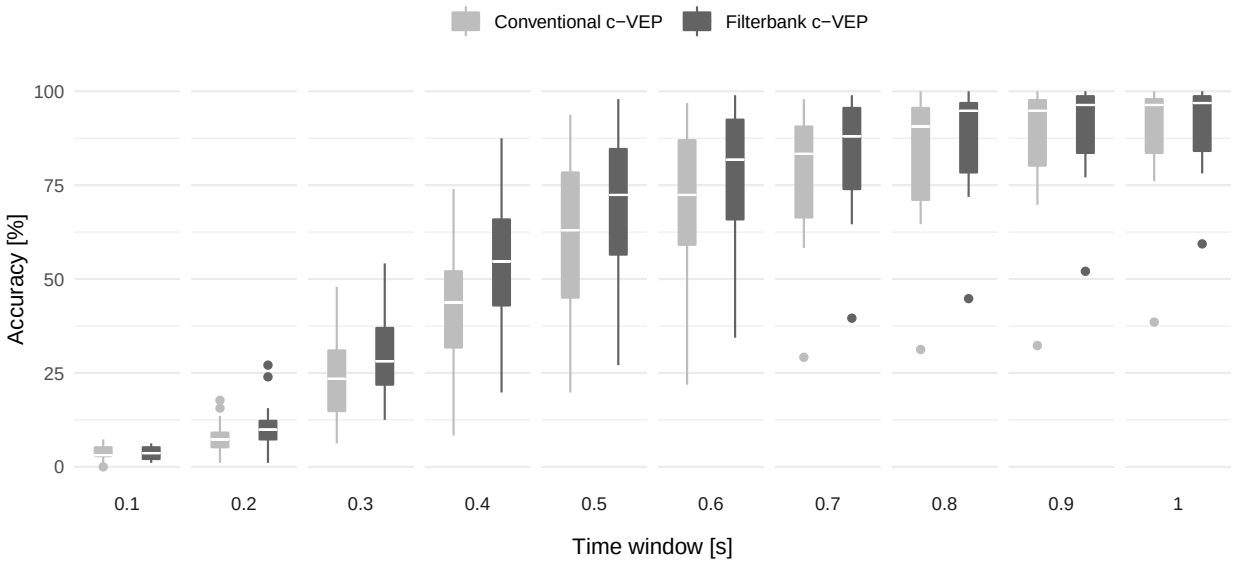



**Figure 5.16:** Accuracies of the conventional c-VEP and filter bank c-VEP. In the box plot, outliers (data points outside 1.5 times the interquartile range) are located outside the "whiskers".

used for training. To investigate the performance boost resulting from the filter bank approach, averaged accuracies for different classification time windows were calculated. Figure 5.16 shows the off-line accuracies for time windows up to 1 s for the conventional and for the filter bank approach. For all evaluated time windows, the median accuracies were higher when applying the filter band classification. Figure 5.17 A shows the individual accuracies separated by the different bands for the 1 s time window. It can be observed that for some subjects, the filter bank approach resulted still in much higher accuracy values in comparison to the standard approach, where only one band is evaluated for classification. For S2, the accuracy increased by more than 20%. It can be seen that for this particular subject, the accuracy and certainty increased when the alpha-band is filtered. Figure 5.17 B shows the averaged off-line certainty level for the 1 s time window, which might reflect on-line performance better. The gamma-band contributed negatively to the certainty level for S2, S11, S15, and S17. For these participants, a filter band design with only two bands might be the better option. Conversely, for S13, the gamma-band showed the highest positive contribution to the certainty value.

Figure 5.18 shows the individual power spectrum density (PSD) estimates, which reflect the variability across subjects as well. For example, S2 showed comparably high alpha activity, most probably unrelated to the stimulus, as filtering the alpha activity yielded higher accuracy according to Figure 5.17. In fact, for this particular subject, reliable control might not have been possible with the standard approach.

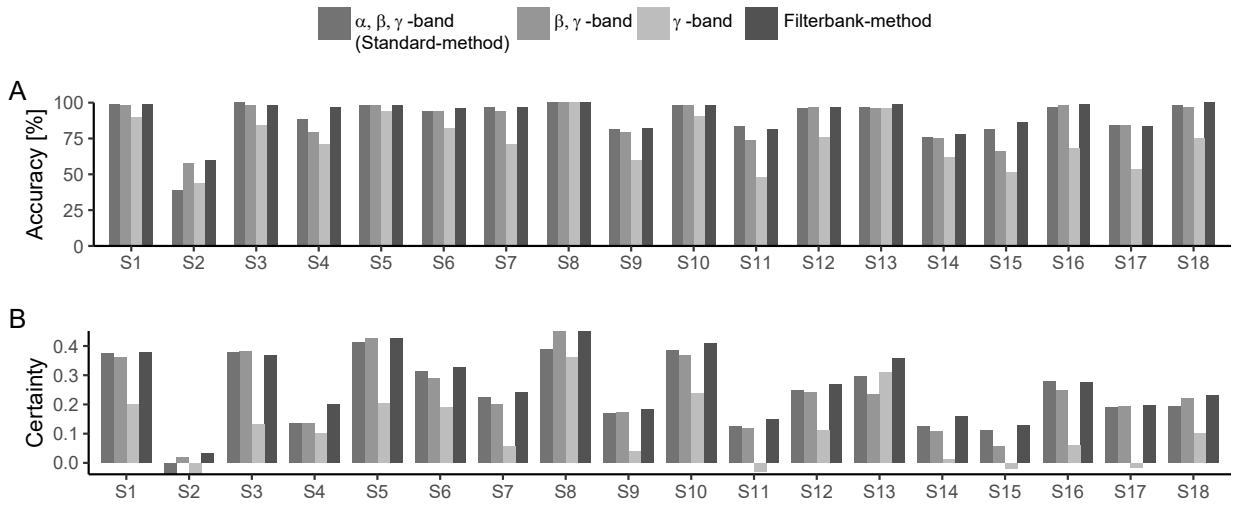



**Figure 5.17:** Off-line accuracies and certainties for different filter banks for a time window of 1s. (A) The accuracies for each subject, averaged over the trials, are shown. (B) The certainty level (the correlation of the target minus the maximum correlation of the remaining classes), averaged over the trials, is shown.

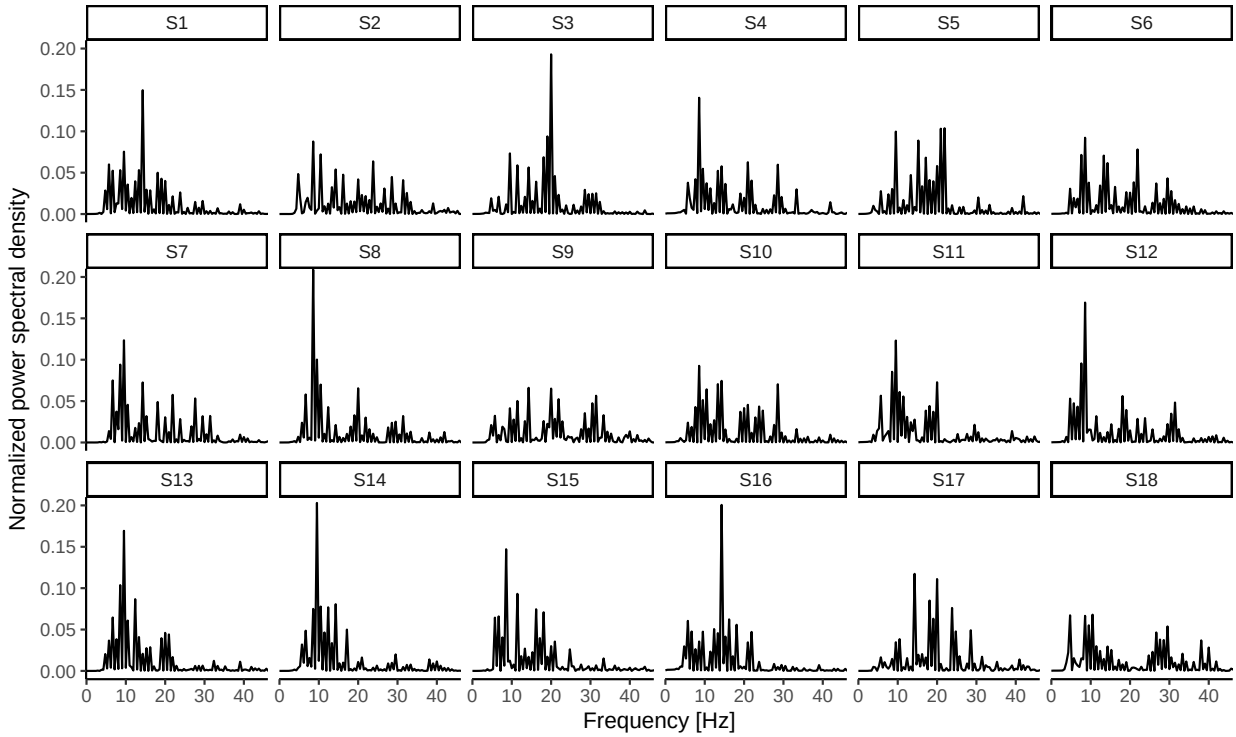



**Figure 5.18:** Normalized power spectrum density (PSD) estimates (sum of the spectrum is normalized to one) of the reference templates at signal channel Oz for all subjects.

**Table 5.7:** On-line results of the c-VEP QWERTZ speller. Provided are the results for letter-by-letter spelling task (BRAIN) and the subject-specific sentence task as listed in Table 5.6 (column Sent.). The individual information transfer rates (ITRs), accuracies (ACCs), and output characters per minute (OCM) are listed.

| | Letter-by-letter | | | Sentence | | |
|---|---|---|---|---|---|---|
| Subject # | ACC [%] | ITR [bpm] | OCM [char/min] | ACC [%] | ITR [bpm] | OCM [char/min] |
| 1 | 100 | 201.3 | 40.3 | 100 | 150.0 | 40.0 |
| 2 | 100 | 65.4 | 13.1 | 94 | 35.1 | 12.2 |
| 3 | 100 | 144.9 | 29.0 | 100 | 68.6 | 18.9 |
| 4 | 100 | 182.9 | 36.6 | 92 | 114.3 | 52.5 |
| 5 | 100 | 229.0 | 45.8 | 100 | 118.3 | 34.9 |
| 6 | 100 | 113.2 | 22.6 | 93 | 104.7 | 45.3 |
| 7 | 100 | 103.5 | 20.7 | 100 | 76.7 | 29.7 |
| 8 | 100 | 245.9 | 49.2 | 94 | 106.2 | 29.0 |
| 9 | 100 | 145.6 | 29.1 | 100 | 111.5 | 49.1 |
| 10 | 100 | 205.5 | 41.1 | 100 | 165.5 | 49.7 |
| 11 | 100 | 103.5 | 20.7 | 92 | 53.0 | 17.5 |
| 12 | 100 | 162.2 | 32.4 | 92 | 32.6 | 14.2 |
| 13 | 100 | 184.1 | 36.8 | 100 | 106.3 | 35.4 |
| 14 | 100 | 161.3 | 32.3 | 95 | 127.9 | 38.0 |
| 15 | 100 | 121.0 | 24.2 | 100 | 87.4 | 30.6 |
| 16 | 100 | 121.0 | 24.2 | 100 | 50.2 | 13.4 |
| 17 | 100 | 86.0 | 17.2 | 94 | 71.6 | 26.9 |
| 18 | 100 | 111.5 | 22.3 | 100 | 96.7 | 39.9 |
| Mean | 100 | 149.3 | 29.9 | 97 | 93.1 | 32.1 |

The on-line spelling tasks were evaluated with classification accuracy, ITR, and OCM. Table 5.7 lists the on-line results for each participant for the letter-by-letter and sentence spelling tasks. For the letter-by-letter spelling task, all participants reached an accuracy of 100%. For the sentence spelling task, accuracies ranged from 92% to 100%. The mean (SD) accuracies were 100% (0) for the letter-by-letter spelling task and 97.0% (3.4) for the sentence spelling task, respectively.

Note that S2, which achieved poor results in the off-line analysis, also reached 100% accuracy for the letter-by-letter task. This is because, in the on-line experiment, the sliding window mechanism was applied; dynamic time windows were used. For S2 the time windows were on average much longer than for the other subjects (the average selection time was 4.5 s for S2). For this reason, the ITR for S2 was the lowest.

The ITR values ranged from 65.4 to 235.9 bpm and from 23.6 to 165.5 bpm for the letter-by-letter and sentence spelling task, respectively. Mean (SD) ITRs of 149.3 (49.3) bpm and 93.1 (36.1) bpm were achieved.

The OCM was slightly higher for the sentence spelling tasks, where participants could use the integrated dictionary feature. The mean (SD) values for the OCM were 29.9 (9.9) char/min for

the letter-by-letter task and 32.1 (12.6) char/min for the sentence task.

### Discussion

In the reported study, an asynchronous multi-target c-VEP application was presented. The system yielded overall high on-line classification accuracies.

Due to the integration of the sliding window mechanism, the variability of accuracy was low, while the variability in the classification times was high across participants. These observations are in contrast to other high-speed state-of-the-art c-VEP spellers, where the times of the spelling tasks are the same across subjects while accuracies vary.

Thanks to the filter bank classification approach, for most users, reliable control was possible with small classification windows below 0.5 s (see Figure 5.16). Because of the high number of classes, high ITRs were achieved. Indeed, the average ITR of 149.3 bpm achieved in the letter-by-letter spelling task is to our knowledge the highest ITR reported in on-line applications using the c-VEP paradigm.

Regarding the signal classification, several improvements can be made. The filter banks were segmented via alpha, beta, and gamma-bands resulting in three separate classifier outputs. The research conducted with SSVEP-based BCIs suggests, however, that a higher number of sub-band components results in better performance (see, e.g., [68]). Moreover, in this study, the correlations corresponding to the different filter banks were averaged. Optimizing weight selection for the different bands could further improve system reliability.

Regarding the $n$-gram based word prediction model, the mean values for the OCM were slightly higher for the sentence spelling task in comparison to the letter-by-letter spelling task. The participants used the interface for the first time; with more training regarding the functionality of the application, the dictionary integration might yield even better performance. The OCM metric is dependent on the complexity of the sentence. Since individual sentences were used, dictionary suggestions were more helpful for participants who had an easier sentence spelling task. For example, subjects 9 and 10 both achieved similar OCM values of roughly 50 char/min. However, the ITRs for these participants were quite different (111.5 and 165.5 bpm). This difference can be explained by the fact that the ITR does not reflect the utilization of the dictionary.

We want to note several limitations. First, the ITR is highly dependent on the pause in between trials (i.e., gaze shifting phase). In preliminary experiments, we explored gaze shifting windows as low as 0.5 s. Some users were able to spell short words (e.g., BCI) with ITRs around 300 bpm using this setup in preliminary test runs. However, for longer sentences, this approach is not reliable; the time to locate the next desired target is not sufficient, especially for untrained users. Second, the study was conducted with young and healthy participants only; their mean age does not reflect the general population. As reported in section 5.2, subject age is correlated with decreased BCI performance. Therefore, especially when considering the target population for spelling applications, the focus should lie on accuracy and a high literacy rate rather than ITR. In preliminary tests with a prototype version of the interface, some healthy users were unable to achieve accuracies above 70% with the presented system. With the eight-target version of the speller, as presented in the previous sections, this issue was not observed.

**SSVEP vs c-VEP for multi-target BCIs**

Multi-target systems can be realized with both the c-VEP and SSVEP paradigm. For the SSVEP paradigm, if sine and cosine reference templates are used, no training session for template recording is necessary. However, the accuracy and literacy rate can be reduced in this case (see section 4.5). A hybrid frequency and phase coding approach can improve performance significantly [114, 155]. When using this method, a training session is also required (see, e.g., [52]). In contrast to the c-VEP paradigm, the trials recorded in the training phase can not be shifted to one single class to generate averages. Because each target carries unique frequency information, the templates need to be generated independently for each frequency class. For a multi-target system, this can lead to much longer training time in comparison to the c-VEP paradigm. Moreover, a sliding window approach, as presented here, is harder to realize, as the common repeating cycle of the targets depends on the frequency resolution. For example, in section 5.3, where only eight classes were used, the difference between targets was 1 Hz; the repeating cycle was 1 s. For multi-target SSVEP systems, a higher resolution needs to be used. Most typically, a 0.2 Hz difference between targets is employed (e.g., 8 Hz, 8.2 Hz, 8.4 Hz, and so on). In this case, the repeating cycle is 5 s, which is rather long for the sliding window approach, where data are shuffled out dynamically. For c-VEP systems, the duration of the repeating cycle is depended on the code length and the refresh rate. Therefore, in comparison to the SSVEP paradigm, more suitable repeating cycles can be applied (in this study, 1.05 s).

For these reasons, the c-VEP application could be the better choice for asynchronous multi-target BCI applications. Due to the high classification accuracies and selection speed, combinations with eye-tracking devices seem to be a logical next step.

# Summary and Future work 6

## 6.1 Summary

This thesis investigated parameter optimization for VEP-based BCIs. BCIs translate recorded brain signals into computer commands. Chapter 2 described, among other BCI approaches, systems based on VEPs, where brain signals are elicited when gazing at a visual stimulus. VEP-based BCIs enable a communication channel for people with severe motor disabilities. Researchers developed various kinds of BCI applications, employing frequency-modulated stimuli (SSVEP paradigm) or code-modulated stimuli (c-VEP paradigm). The human user is the most variable factor in the BCI framework due to the complexity of the individual brain activity. A primary goal of this work was to reduce the BCI illiteracy rate, which is the percentage of users who are not able to gain control over the system. A focus was laid on asynchronous spelling applications as these handle inter- and within-subject better than synchronous systems which rely on fixed time windows for classification.

Chapter 3 described implementations of the two major VEP strategies, namely the SSVEP and the c-VEP paradigm. For the SSVEP systems, the frequency-based flickering can be generated using various kinds of modulation techniques. The frequency selection is, however, limited by the vertical refresh rate. For the c-VEP systems, the flickering pattern is determined by the vertical refresh rate and the code length. As c-VEP systems rely on phase information, data acquisition, and stimulus presentation need to be synchronized. For this synchronization, a software-based approach was presented.

Chapter 4 investigated SSVEP-BCI parameters and their influence on BCI spelling performance. The investigated variables included the stimuli selection, length of the classification time window, the flickering speed, the user age, and the number of targets.

Section 4.3 investigated to what extend user age impacts BCI performance. The reported study revealed that older adults have slightly poorer control over the system. The mean ITR of the young age group was 27.4 bpm and 16.1 bpm for the elderly group. The results indicate that the user age should be considered when designing an SSVEP-based application. The average classification time window length, a key parameter for the BCI, was usually larger for the participants of advanced age.

Section 4.4 presented an automated calibration software, which was tested with 61 participants. Unfortunately, the achieved mean ITR of 21.9 bpm was below the values for VEP systems that can be found in the literature. One way to yield higher ITRs, in theory, is to increase the number of classes.

Section 4.5 investigated the possibilities and limitations in regard to the number of targets for SSVEP-systems. Unfortunately, the reported studies showed a negative correlation between the number of targets and the classification accuracy and BCI literacy rate.

Another way to improve system speed is the incorporation of personalized EEG data. For this, a training session needs to be conducted, where several trials of EEG data are recorded. This is a standard approach for the c-VEP paradigm and can also be used for the SSVEP paradigm. A disadvantage of c-VEP based systems is that they are typically implemented as synchronous systems (i.e., after fixed preset time periods, the system produces command outputs); users have only a limited time to locate the desired target. Chapter 5 focused on c-VEP based systems. Section 5.1 presented an asynchronous c-VEP implementation using threshold-based target identification and a dynamic time window mechanism. These methods were tested with a dictionary-driven spelling application using eight flashing targets. The GUI presented three word suggestions that were updated after each letter selection. While it achieved lower ITRs than other c-VEP state-of-the-art spellers, in terms of character output speed, the application could compete. Section 5.2 investigated different flickering speeds and age-related performance differences using this interface. The elderly participants achieved poorer BCI control again; the performance difference between young and elderly users was most striking for the fastest flickering speed, generated with a refresh rate of 120 Hz.

Section 5.3 explored personalized dynamic classification time windows and thresholds for the proposed asynchronous system. The optimization techniques were not only tested with the c-VEP paradigm but also with the frequency and phase coded SSVEP stimulus design. The spelling performance of twelve healthy participants was evaluated. All participants completed sentence spelling tasks, reaching high average accuracies of 94% and 96.3% for the c-VEP and the SSVEP paradigm, respectively. Average ITRs around 57 bpm were achieved for both paradigms. Questionnaire results indicate that the c-VEP flickering was perceived as slightly more annoying.

Finally, section 5.4 presented further improvements of the c-VEP system, with regard to the signal classification algorithms. Moreover, the number of targets was increased to 32; the sliding window mechanism and the dictionary integration were transferred to a

single-step QWERTZ style application. In the reported study, 18 participants yielded high mean accuracies in letter-by-letter and sentence spelling, (100% and 97%, respectively). The corresponding mean ITRs were 149.3 bpm and 93.1 bpm. Moreover, the literacy rate and accuracy remained high when the number of c-VEP targets was increased from 8 to 32. If one compares the applications which use sine and cosine reference templates (chapter 4) to the applications using personalized EEG data (chapter 5), the following can be observed: Using personalized EEG data can increase overall system performance as a much shorter classification time window can be used in asynchronous applications. The data obtained from the recording session for the EEG templates can also be used to optimize parameters for asynchronous applications.

## 6.2 Conclusion

This research aimed to identify system optimization strategies for VEP-based BCIs. One of the general goals was to investigate age-related differences in BCI performance. Based on comparative studies with different age groups, it can be concluded that user age is indeed an essential factor to consider when designing the BCI. The results of the reported study [148] indicate that the performance gap between elderly and young users increases with the flickering rate.

This work was further set out to examine key system parameters and their impact on BCI performance and BCI illiteracy. In chapter 4, parameters of training free SSVEP systems were investigated. Sinusoidal reference signals were used to identify the stimulation frequency at which the user gazed. In contrast, in chapter 5, personalized reference signals were recorded in a training session. The overall literacy rate, accuracy, and ITR were higher using the latter approach, which takes the inter-user variability of VEP responses into account. This is in line with the findings of Zerafa et al. [114], according to which training-free systems are outperformed by systems incorporating user-specific training data.

Another primary goal of the thesis was to improve the VEP signal classification. In terms of spatial filtering, the well-established methods CCA and MEC were used. For the c-VEP paradigm, we adopted an ensemble-based approach, where multiple spatial filters depending on the number of stimulus classes are used. This approach yielded better results than the conventional approach [72]. In the frequency domain, the filter bank approach, introduced by Chen et al. [68] for the SSVEP paradigm was applied to the c-VEP paradigm resulting in significantly faster performance according to our off-line analysis [150].

Since our main aim was, as mentioned in the introduction, to implement auto-calibration methods, we developed a wizard for SSVEP-based BCIs that determines the optimal time window among other parameters on the basis of a small calibration session [86]. Moreover, a modified calibration software has been developed for the training-based systems (c-VEP and frequency and phase coded SSVEP-BCIs) [149].

The above findings and observations were synthesized to implement a robust BCI spelling application that can be set up by non-experts [72, 149]. Due to the developed dynamic time window mechanism, the proposed system can be characterized as asynchronous BCI. If the classification threshold is set carefully, the BCI can distinguish between intentional and unintentional fixation. This is an essential feature in terms of usability of spelling applications, as it allows for more complex interfaces. In this respect, an $n$-gram based word prediction model was integrated. Eight-target and 32-target implementations of the system were tested; the results were quite promising. The 32-target interface yielded fast performance in a test with 18 healthy participants [150]; The eight-target interface has been tested in total with more than 50 users including 13 elderly participants [72, 89, 148, 149]. The developed application was at the time of writing also being tested in ongoing preliminary experiments with disabled participants. These experiments were conducted by students of the University of Cologne. While the reported accuracies were generally much lower, successful tests with the developed GUI were reported with participants suffering from ALS and cerebral palsy.

## 6.3 Future Work

In the following, avenues for future research are outlined:

**Further optimization of the wizard software:**
The setup of individual parameters was mainly focused on parameters relevant for asynchronous on-line applications. On the basis of off-line data, several other parameters could be optimized: the lower and upper cutoff frequencies for the filter bank design, the weights for the different filter banks, and the number of weights used for the design of spatial filters.

**Classification with neural networks:**
Small training sets were collected to reduce the duration of the experiment. In most of the studies reported in chapter 5, training data for each class were averaged over six trials. Better performance might be achievable with longer training

sessions. With larger data sets, deep neural networks might yield better performance while simplifying feature extraction.

**Evaluation of the effectiveness of continuous BCI feedback:**
The type of continuous feedback given to the user could have an impact on the overall performance. In this thesis, continuous feedback was provided in the form of progress bars or by varying the size of the targets. Different ideas to provide feedback based on the current classifier state could be investigated and evaluated.

**More subtle visual stimuli for the c-VEP paradigm:**
According to the conducted questionnaires, the flickering is perceived as annoying by many users, especially for the c-VEP paradigm where the stimulus design is realized as a black and white pattern on the basis of a binary $m$-sequence. Instead, one could employ ternary (base 3) or quinary (base 5) $m$-sequences for stimulus design, which have good autocorrelation as well. Instead of a black-white pattern, different gray shades could be used, yielding a more subtle stimulation.

**VEP/eye-tracking hybrid:**
Eye-tracking devices are communications tools for disabled people who have still oculomotor control. Complications such as accidental selections, slow dwell-based classification, and low accuracy could be improved by combining it with a BCI. The asynchronous VEP applications described in this thesis are not affected by the Midas touch problem.

**Tests with disabled users:**
Research indicates that brain responses of disabled participants are harder to interpret, resulting in lower accuracies and literacy rates. For this population, customization based on high accuracy rather than system speed could be a better option. The presented application was at the time of writing being tested with patients.

**Improved dictionary implementation:**
The word suggestion mechanism can be improved. For example, automatic error correction could be implemented into the dictionary-driven spelling interface.

**Investigate recalibration for personalized EEG data:**
In the experiments reported in this thesis, EEG data to generate templates was recorded immediately before the copy spelling phases. In long term use, recalibration might be necessary, depending on the quality of the calibration. During development and testing, we found that if electrode placement is the same, EEG data collected several months prior to the test still yielded proper calibration. Future research could

investigate the longevity of recorded EEG data in terms of on-line BCI applications.

# Bibliography


[1] I. Volosyak, H. Cecotti, D. Valbuena, and A. Gräser. 'Evaluation of the Bremen SSVEP Based BCI in Real World Conditions'. In: *2009 IEEE International Conference on Rehabilitation Robotics*. The Community (ICORR). Kyoto, Japan, June 2009, pp. 322–331. DOI: 10.1109/ICORR.2009.5209543.

[2] G. Müller-Putz and G. Pfurtscheller. 'Control of an Electrical Prosthesis With an SSVEP-Based BCI'. In: *IEEE Transactions on Biomedical Engineering* 55.1 (Jan. 2008), pp. 361–364. DOI: 10.1109/TBME.2007.897815.

[3] P. Martinez, H. Bakardjian, and A. Cichocki. 'Fully Online Multicommand Brain-Computer Interface with Visual Neurofeedback Using SSVEP Paradigm'. In: *Computational Intelligence and Neuroscience* 2007 (2007), pp. 1–9. DOI: 10.1155/2007/94561.

[4] D. E. Thompson, L. R. Quitadamo, L. Mainardi, K. U. R. Laghari, S. Gao, P.-J. Kindermans, J. D. Simeral, R. Fazel-Rezai, M. Matteucci, T. H. Falk, L. Bianchi, C. A. Chestek, and J. E. Huggins. 'Performance Measurement for Brain–Computer or Brain–Machine Interfaces: A Tutorial'. In: *Journal of Neural Engineering* 11.3 (June 1, 2014), p. 035001. DOI: 10.1088/1741-2560/11/3/035001.

[5] S. M. T. Müller, P. F. Diez, T. F. Bastos-Filho, M. Sarcinelli-Filho, V. Mut, E. Laciar, and E. Avila. 'Robotic Wheelchair Commanded by People with Disabilities Using Low/High-Frequency SSVEP-Based BCI'. In: *World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada*. Ed. by D. A. Jaffray. Vol. 51. Cham: Springer International Publishing, 2015, pp. 1177–1180. DOI: 10.1007/978-3-319-19387-8_285.

[6] I. Volosyak, H. Cecotti, and A. Gräser. 'Impact of Frequency Selection on LCD Screens for SSVEP Based Brain-Computer Interfaces'. In: *Bio-Inspired Systems: Computational and Ambient Intelligence*. Ed. by J. Cabestany, F. Sandoval, A. Prieto, and J. M. Corchado. Vol. 5517. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 706–713. DOI: 10.1007/978-3-642-02478-8_88.

[7] E. E. Sutter. 'The Brain Response Interface: Communication through Visually-Induced Electrical Brain Responses'. In: *Journal of Microcomputer Applications* 15.1 (Jan. 1992), pp. 31–45. DOI: 10.1016/0745-7138(92)90045-7.

[8] I. Volosyak, D. Valbuena, T. Luth, T. Malechka, and A. Gräser. 'BCI Demographics II: How Many (and What Kinds of) People Can Use a High-Frequency SSVEP BCI?' In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 19.3 (June 2011), pp. 232–239. DOI: 10.1109/TNSRE.2011.2121919.

[9] M. Spüler, W. Rosenstiel, and M. Bogdan. 'Online Adaptation of a C-VEP Brain-Computer Interface(BCI) Based on Error-Related Potentials and Unsupervised Learning'. In: *PLoS ONE* 7.12 (Dec. 7, 2012). Ed. by M. Baumert, e51077. DOI: 10.1371/journal.pone.0051077.

[10] J. J. Vidal. 'Toward Direct Brain-Computer Communication'. In: *Annual review of Biophysics and Bioengineering* 2.1 (1973), pp. 157–180. DOI: 10.1146/annurev.bb.02.060173.001105.

[11] J. J. Vidal. 'Real-Time Detection of Brain Events in EEG'. In: *Proceedings of the IEEE* 65.5 (1977), pp. 633–641. DOI: 10.1109/PROC.1977.10542.

[12] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. 'Brain–Computer Interfaces for Communication and Control'. In: *Clinical Neurophysiology* 113.6 (June 1, 2002), pp. 767–791. DOI: 10.1016/S1388-2457(02)00057-3.

[13] A. Rezeika, M. Benda, P. Stawicki, F. Gembler, A. Saboor, and I. Volosyak. 'Brain-Computer Interface Spellers: A Review'. In: *Brain Sciences* 8.4 (2018). DOI: 10.3390/brainsci8040057.

[14] P. Majaranta and A. Bulling. 'Eye Tracking and Eye-Based Human–Computer Interaction'. In: *Advances in Physiological Computing*. Ed. by S. H. Fairclough and K. Gilleade. London: Springer London, 2014, pp. 39–65. DOI: 10.1007/978-1-4471-6392-3_3.

[15] R. Sharma, S. Hicks, C. M. Berna, C. Kennard, K. Talbot, and M. R. Turner. 'Oculomotor Dysfunction in Amyotrophic Lateral Sclerosis: A Comprehensive Review'. In: *Archives of neurology* 68.7 (2011), pp. 857–861. DOI: 10.1001/archneurol.2011.130.

[16] L. F. Nicolas-Alonso and J. Gomez-Gil. 'Brain Computer Interfaces, a Review'. In: *Sensors* 12.2 (Jan. 31, 2012), pp. 1211–1279. DOI: 10.3390/s120201211.

[17] H. Berger. 'Über das Elektrenkephalogramm des Menschen'. In: *European archives of psychiatry and clinical neuroscience* 87.1 (1929), pp. 527–570. DOI: https://doi.org/10.1007/BF01797193.

[18] M. Teplan. 'Fundamentals of EEG Measurement'. In: *Measurement Science Review* 2.2 (2002), pp. 1–11.

[19] R. Oostenveld and P. Praamstra. 'The Five Percent Electrode System for High-Resolution EEG and ERP Measurements'. In: *Clinical Neurophysiology* 112.4 (Apr. 2001), pp. 713–719. DOI: 10.1016/S1388-2457(00)00527-7.

[20] B. Burle, L. Spieser, C. Roger, L. Casini, T. Hasbroucq, and F. Vidal. 'Spatial and Temporal Resolutions of EEG: Is It Really Black and White? A Scalp Current Density View'. In: *International Journal of Psychophysiology* 97.3 (Sept. 2015), pp. 210–220. DOI: 10.1016/j.ijpsycho.2015.05.004.

[21] G. A. Light, L. E. Williams, F. Minow, J. Sprock, A. Rissling, R. Sharp, N. R. Swerdlow, and D. L. Braff. 'Electroencephalography (EEG) and Event-Related Potentials (ERPs) with Human Participants'. In: *Current Protocols in Neuroscience* 52.1 (July 2010), pp. 6.25.1–6.25.24. DOI: 10.1002/0471142301.ns0625s52.

[22] H. Jasper. 'The Ten-Twenty Electrode System of the International Federation'. In: *Electroencephalogr. Clin. Neurophysiol.* 10 (1958), pp. 371–375.

[23] I. Volosyak, D. Valbuena, T. Malechka, J. Peuscher, and A. Gräser. 'Brain–Computer Interface Using Water-Based Electrodes'. In: *Journal of Neural Engineering* 7.6 (Dec. 1, 2010), p. 066007. DOI: 10.1088/1741-2560/7/6/066007.

[24] V. Mihajlović, G. Garcia-Molina, and J. Peuscher. 'Dry and Water-Based EEG Electrodes in SSVEP-Based BCI Applications'. In: *Biomedical Engineering Systems and Technologies*. Ed. by J. Gabriel, J. Schier, S. Van Huffel, E. Conchon, C. Correia, A. Fred, and H. Gamboa. Vol. 357. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 23–40. DOI: 10.1007/978-3-642-38256-7_2.

[25] M. Spüler. 'A High-Speed Brain-Computer Interface (BCI) Using Dry EEG Electrodes'. In: *PLOS ONE* 12.2 (Feb. 22, 2017), e0172400. DOI: 10.1371/journal.pone.0172400.

[26] D. Regan. *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. New York: Elsevier, 1989.

[27] B. Blankertz, M. Krauledat, G. Dornhege, J. Williamson, R. Murray-Smith, and K.-R. Müller. 'A Note on Brain Actuated Spelling with the Berlin Brain-Computer Interface'. In: *International Conference on Universal Access in Human-Computer Interaction*. Springer, 2007, pp. 759–768.

[28] L. A. Farwell and E. Donchin. 'Talking off the Top of Your Head: Toward a Mental Prosthesis Utilizing Event-Related Brain Potentials'. In: *Electroencephalography and clinical Neurophysiology* 70.6 (1988), pp. 510–523. DOI: 10.1016/0013-4694(88)90149-6.

[29] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K.-R. Müller, G. Curio, and T. Dickhaus. 'Neurophysiological Predictor of SMR-Based BCI Performance'. In: *NeuroImage* 51.4 (July 2010), pp. 1303–1309. DOI: 10.1016/j.neuroimage.2010.03.022.

[30] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer. 'EEG-Based Discrimination between Imagination of Right and Left Hand Movement'. In: *Electroencephalography and Clinical Neurophysiology* 103.6 (Dec. 1997), pp. 642–651. DOI: 10.1016/S0013-4694(97)00080-1.

[31] D. J. McFarland, L. A. Miner, T. M. Vaughan, and J. R. Wolpaw. 'Mu and Beta Rhythm Topographies during Motor Imagery and Actual Movements'. In: *Brain topography* 12.3 (2000), pp. 177–186. DOI: 10.1023/A:1023437823106.

[32] H. Jasper and W. Penfield. 'Electrocorticograms in Man: Effect of Voluntary Movement upon the Electrical Activity of the Precentral Gyrus'. In: *Archiv für Psychiatrie und Nervenkrankheiten* 183.1-2 (1949), pp. 163–174. DOI: 10.1007/BF01062488.

[33] A. Kübler, F. Nijboer, J. Mellinger, T. M. Vaughan, H. Pawelzik, G. Schalk, D. J. McFarland, N. Birbaumer, and J. R. Wolpaw. 'Patients with ALS Can Use Sensorimotor Rhythms to Operate a Brain-Computer Interface'. In: *Neurology* 64.10 (May 24, 2005), pp. 1775–1777. DOI: 10.1212/01.WNL.0000158616.43002.6D.

[34] G. Pfurtscheller and C. Neuper. 'Motor Imagery and Direct Brain-Computer Communication'. In: *Proceedings of the IEEE* 89.7 (2001), pp. 1123–1134. DOI: 10.1109/5.939829.

[35] M. Krauledat, M. Tangermann, B. Blankertz, and K.-R. Müller. 'Towards Zero Training for Brain-Computer Interfacing'. In: *PLoS ONE* 3.8 (Aug. 13, 2008), e2967. DOI: 10.1371/journal.pone.0002967.

[36] S. Sutton, M. Braren, J. Zubin, and E. R. John. 'Evoked-Potential Correlates of Stimulus Uncertainty'. In: *Science* 150.3700 (Nov. 26, 1965), pp. 1187–1188. DOI: 10.1126/science.150.3700.1187.

[37] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger. 'How Many People Are Able to Control a P300-Based Brain-Computer Interface (BCI)?' In: *Neuroscience Letters* 462.1 (Sept. 2009), pp. 94–98. DOI: 10.1016/j.neulet.2009.06.045.

[38] P. Yuan, X. Gao, B. Allison, Y. Wang, G. Bin, and S. Gao. 'A Study of the Existing Problems of Estimating the Information Transfer Rate in Online Brain-Computer Interfaces'. In: *Journal of Neural Engineering* 10.2 (Apr. 1, 2013), p. 026014. DOI: 10.1088/1741-2560/10/2/026014.

[39] F. Nijboer, E. Sellers, J. Mellinger, M. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. Krusienski, T. Vaughan, J. Wolpaw, N. Birbaumer, and A. Kübler. 'A P300-Based Brain–Computer Interface for People with Amyotrophic Lateral Sclerosis'. In: *Clinical Neurophysiology* 119.8 (Aug. 2008), pp. 1909–1916. DOI: 10.1016/j.clinph.2008.03.034.

[40] X. Gao, X. Xu, M. Cheng, and S. Gao. 'A BCI-Based Environmental Controller for the Motion-Disabled'. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 11.2 (June 2003), pp. 137–140. DOI: 10.1109/TNSRE.2003.814449.

[41] F.-B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki. 'Steady-State Visually Evoked Potentials: Focus on Essential Paradigms and Future Perspectives'. In: *Progress in Neurobiology* 90.4 (Apr. 2010), pp. 418–438. DOI: 10.1016/j.pneurobio.2009.11.005.

[42] I. Volosyak. 'SSVEP-Based Bremen–BCI Interface—Boosting Information Transfer Rates'. In: *Journal of Neural Engineering* 8.3 (June 1, 2011), p. 036020. DOI: 10.1088/1741-2560/8/3/036020.

[43] C. S. Herrmann. 'Human EEG Responses to 1-100 Hz Flicker: Resonance Phenomena in Visual Cortex and Their Potential Correlation to Cognitive Phenomena'. In: *Experimental Brain Research* 137.3-4 (Apr. 2, 2001), pp. 346–353. DOI: 10.1007/s002210100682.

[44] E. E. Sutter. 'The Visual Evoked Response as a Communication Channel'. In: *Proceedings on IEEE 1984 Symposium on Biosensors*. 1984, pp. 95–100.

[45] G. Bin, X. Gao, Y. Wang, B. Hong, and S. Gao. 'VEP-Based Brain-Computer Interfaces: Time, Frequency, and Code Modulations [Research Frontier]'. In: *IEEE Computational Intelligence Magazine* 4.4 (Nov. 2009), pp. 22–26. DOI: 10.1109/MCI.2009.934562.

[46] G. Bin, X. Gao, Y. Wang, Y. Li, B. Hong, and S. Gao. 'A High-Speed BCI Based on Code Modulation VEP'. In: *Journal of Neural Engineering* 8.2 (Mar. 2011), p. 025015. DOI: 10.1088/1741-2560/8/2/025015.

[47] S. Kelly, E. Lalor, R. Reilly, and J. Foxe. 'Visual Spatial Attention Tracking Using High-Density SSVEP Data for Independent Brain-Computer Communication'. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 13.2 (June 2005), pp. 172–178. DOI: 10.1109/TNSRE.2005.847369.

[48] P.-L. Lee, J.-C. Hsieh, C.-H. Wu, K.-K. Shyu, S.-S. Chen, T.-C. Yeh, and Y.-T. Wu. 'The Brain Computer Interface Using Flash Visual Evoked Potential and Independent Component Analysis'. In: *Annals of Biomedical Engineering* 34.10 (Oct. 23, 2006), pp. 1641–1654. DOI: 10.1007/s10439-006-9175-8.

[49] W. Skrandies, A. Jedynak, and R. Kleiser. 'Scalp Distribution Components of Brain Activity Evoked by Visual Motion Stimuli'. In: *Experimental brain research* 122.1 (1998), pp. 62–70. DOI: 10.1007/s002210050491.

[50] X. Zhang, G. Xu, J. Xie, and X. Zhang. 'Brain Response to Luminance-Based and Motion-Based Stimulation Using Inter-Modulation Frequencies'. In: *PLOS ONE* 12.11 (Nov. 15, 2017). Ed. by H. He, e0188073. DOI: 10.1371/journal.pone.0188073.

[51] S. Mason and G. Birch. 'A General Framework for Brain-Computer Interface Design'. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 11.1 (Mar. 2003), pp. 70–85. DOI: 10.1109/TNSRE.2003.810426.

[52] M. Nakanishi, Y. Wang, X. Chen, Y.-T. Wang, X. Gao, and T.-P. Jung. 'Enhancing Detection of SSVEPs for a High-Speed Brain Speller Using Task-Related Component Analysis'. In: *IEEE Transactions on Biomedical Engineering* 65.1 (Jan. 2018), pp. 104–112. DOI: 10.1109/TBME.2017.2694818.

[53] J. Thielen, P. van den Broek, J. Farquhar, and P. Desain. 'Broad-Band Visually Evoked Potentials: Re(Con)Volution in Brain-Computer Interfacing'. In: *PLOS ONE* 10.7 (July 24, 2015). Ed. by D. Marinazzo, e0133797. DOI: 10.1371/journal.pone.0133797.

[54] M. Spüler, W. Rosenstiel, and M. Bogdan. 'One Class SVM and Canonical Correlation Analysis Increase Performance in a C-VEP Based Brain-Computer Interface (BCI).' In: *ESANN*. 2012. DOI: 10.13140/2.1.2186.7526.

[55] T.-H. Nguyen and W.-Y. Chung. 'A Single-Channel SSVEP-Based BCI Speller Using Deep Learning'. In: *IEEE Access* 7 (2019), pp. 1752–1763. DOI: 10.1109/ACCESS.2018.2886759.

[56] C.-T. Lin, C.-J. Chang, B.-S. Lin, S.-H. Hung, C.-F. Chao, and I.-J. Wang. 'A Real-Time Wireless Brain–Computer Interface System for Drowsiness Detection'. In: *IEEE Transactions on Biomedical Circuits and Systems* 4.4 (Aug. 2010), pp. 214–222. DOI: 10.1109/TBCAS.2010.2046415.

[57] B. Blankertz, L. Acqualagna, S. Dähne, S. Haufe, M. Schultze-Kraft, I. Sturm, M. Uščumlic, M. A. Wenzel, G. Curio, and K.-R. Müller. 'The Berlin Brain-Computer Interface: Progress Beyond Communication and Control'. In: *Frontiers in Neuroscience* 10 (Nov. 21, 2016). DOI: 10.3389/fnins.2016.00530.

[58] H. Riechmann, A. Finke, and H. Ritter. 'Using a cVEP-Based Brain-Computer Interface to Control a Virtual Agent'. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 24.6 (June 2016), pp. 692–699. DOI: 10.1109/TNSRE.2015.2490621.

[59] M. Adams, M. Benda, A. Saboor, A. F. Krause, A. Rezeika, F. Gembler, P. Stawicki, M. Hesse, K. Essig, S. Ben-Salem, S. Islam, A. Vogelsang, T. Jungeblut, U. Rückert, and I. Volosyak. 'Towards an SSVEP-BCI Controlled Smart Home'. In: *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). Oct. 2019, pp. 2737–2742. DOI: 10.1109/SMC.2019.8914668.

[60] S. R. Soekadar, N. Birbaumer, M. W. Slutzky, and L. G. Cohen. 'Brain–Machine Interfaces in Neurorehabilitation of Stroke'. In: *Neurobiology of Disease* 83 (Nov. 2015), pp. 172–179. DOI: 10.1016/j.nbd.2014.11.025.

[61] F. Gembler, P. Stawicki, and I. Volosyak. 'Suitable Number of Visual Stimuli for SSVEP-Based BCI Spelling Applications'. In: *Advances in Computational Intelligence: 14th International Work-Conference on Artificial Neural Networks, IWANN 2017, Cadiz, Spain, June 14-16, 2017, Proceedings, Part II*. Ed. by I. Rojas, G. Joya, and A. Catala. Cham: Springer International Publishing, 2017, pp. 441–452. DOI: 10.1007/978-3-319-59147-6_38.

[62] F. Gembler, P. Stawicki, and I. Volosyak. 'A Comparison of SSVEP-Based BCI-Performance Between Different Age Groups'. In: *Advances in Computational Intelligence: 13th International Work-Conference on Artificial Neural Networks, IWANN 2015, Palma de Mallorca, Spain, June 10-12, 2015. Proceedings, Part I*. Ed. by I. Rojas, G. Joya, and A. Catala. Cham: Springer International Publishing, 2015, pp. 71–77. DOI: 10.1007/978-3-319-19258-1_6.

[63] I. Volosyak, F. Gembler, and P. Stawicki. 'Age-Related Differences in SSVEP-Based BCI Performance'. In: *Neurocomputing* 250 (2017), pp. 57–64. DOI: 10.1016/j.neucom.2016.08.121.

[64] H.-J. Hwang, J.-H. Lim, Y.-J. Jung, H. Choi, S. W. Lee, and C.-H. Im. 'Development of an SSVEP-Based BCI Spelling System Adopting a QWERTY-Style LED Keyboard'. In: *Journal of Neuroscience Methods* 208.1 (June 2012), pp. 59–65. DOI: 10.1016/j.jneumeth.2012.04.011.

[65] Y. Wang, Wang, Y.-T., and Jung, T.-P. 'Visual Stimulus Design for High-Rate SSVEP BCI'. In: *Electronics letters* 46.15 (2010), pp. 1057–1058. DOI: 10.1049/el.2010.0923.

[66] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao. 'High-Speed Spelling with a Noninvasive Brain-Computer Interface'. In: *Proceedings of the National Academy of Sciences* 112.44 (Nov. 3, 2015), E6058–E6067. DOI: 10.1073/pnas.1508080112.

[67] P. Stawicki, F. Gembler, A. Rezeika, and I. Volosyak. 'A Novel Hybrid Mental Spelling Application Based on Eye Tracking and SSVEP-Based BCI'. In: *Brain Sciences* 7.4 (Apr. 5, 2017), p. 35. DOI: 10.3390/brainsci7040035.

[68] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao. 'Filter Bank Canonical Correlation Analysis for Implementing a High-Speed SSVEP-Based Brain-Computer Interface'. In: *Journal of neural engineering* 12.046008 (2015), p. 046008. DOI: 10.1088/1741-2560/12/4/046008.

[69] J. N. da Cruz, F. Wan, C. M. Wong, and T. Cao. 'Adaptive Time-Window Length Based on Online Performance Measurement in SSVEP-Based BCIs'. In: *Neurocomputing* 149 (Feb. 2015), pp. 93–99. DOI: 10.1016/j.neucom.2014.01.062.

[70] H. Cecotti. 'A Self-Paced and Calibration-Less SSVEP-Based Brain-Computer Interface Speller'. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18.2 (Apr. 2010), pp. 127–133. DOI: 10.1109/TNSRE.2009.2039594.

[71] J. Pan, Y. Li, R. Zhang, Z. Gu, and F. Li. 'Discrimination Between Control and Idle States in Asynchronous SSVEP-Based Brain Switches: A Pseudo-Key-Based Approach'. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 21.3 (May 2013), pp. 435–443. DOI: 10.1109/TNSRE.2013.2253801.

[72] F. Gembler and I. Volosyak. 'A Novel Dictionary-Driven Mental Spelling Application Based on Code-Modulated Visual Evoked Potentials'. In: *Computers* 8.2 (2019). DOI: 10.3390/computers8020033.

[73] W. Speier, C. Arnold, and N. Pouratian. 'Integrating Language Models into Classifiers for BCI Communication: A Review'. In: *Journal of Neural Engineering* 13.3 (June 1, 2016), p. 031002. DOI: 10.1088/1741-2560/13/3/031002.

[74] D. B. Ryan, G. E. Frye, G. Townsend, D. R. Berry, S. Mesa-G, N. A. Gates, and E. W. Sellers. 'Predictive Spelling With a P300-Based Brain-Computer Interface: Increasing the Rate of Communication'. In: *International Journal of Human-Computer Interaction* 27.1 (Dec. 30, 2010), pp. 69–84. DOI: 10.1080/10447318.2011.535754.

[75] T. Kaufmann, S. Völker, L. Gunesch, and A. Kübler. 'Spelling Is Just a Click Away – A User-Centered Brain-Computer Interface Including Auto-Calibration and Predictive Text Entry'. In: *Frontiers in Neuroscience* 6 (2012). DOI: 10.3389/fnins.2012.00072.

[76] I. Volosyak, A. Moor, and A. Gräser. 'A Dictionary-Driven SSVEP Speller with a Modified Graphical User Interface'. In: *Advances in Computational Intelligence*. Ed. by J. Cabestany, I. Rojas, and G. Joya. Vol. 6691. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 353–361. DOI: 10.1007/978-3-642-21501-8_44.

[77] A. Schlögl, J. Kronegg, J. E. Huggins, and S. G. Mason. English. In: *Toward Brain-Computer Interfacing*. 1st ed. Neural information processing series. MIT Press, 2007, pp. 327–342.

[78] R. Kohavi. 'A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection'. In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*. IJCAI'95. Montreal, Quebec, Canada: Morgan Kaufmann Publishers Inc., 1995, pp. 1137–1143.

[79] C. E. Shannon. 'A Mathematical Theory of Communication'. In: *Bell System Technical Journal* 27 (1948), pp. 379–423.

[80] B. Dal Seno, M. Matteucci, and L. Mainardi. 'The Utility Metric: A Novel Method to Assess the Overall Performance of Discrete Brain-Computer Interfaces'. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18.1 (Feb. 2010), pp. 20–28. DOI: 10.1109/TNSRE.2009.2032642.

[81] W. Speier, C. Arnold, and N. Pouratian. 'Evaluating True BCI Communication Rate through Mutual Information and Language Models'. In: *PLoS ONE* 8.10 (Oct. 22, 2013), e78432. DOI: 10.1371/journal.pone.0078432.

[82] B. Allison, T. Luth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Gräser. 'BCI Demographics: How Many (and What Kinds of) People Can Use an SSVEP BCI?' In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18.2 (Apr. 2010), pp. 107–116. DOI: 10.1109/TNSRE.2009.2039495.

[83] M. C. Thompson. 'Critiquing the Concept of BCI Illiteracy'. In: *Science and Engineering Ethics* 25.4 (Aug. 2019), pp. 1217–1233. DOI: 10.1007/s11948-018-0061-1.

[84] J. Perelmouter and N. Birbaumer. 'A Binary Spelling Interface with Random Errors'. In: *IEEE Transactions on Rehabilitation Engineering* 8.2 (June 2000), pp. 227–232. DOI: 10.1109/86.847824.

[85] C. Brunner, B. Z. Allison, D. J. Krusienski, V. Kaiser, G. R. Müller-Putz, G. Pfurtscheller, and C. Neuper. 'Improved Signal Processing Approaches in an Offline Simulation of a Hybrid Brain–Computer Interface'. In: *Journal of Neuroscience Methods* 188.1 (Apr. 2010), pp. 165–173. DOI: 10.1016/j.jneumeth.2010.02.002.

[86] F. Gembler, P. Stawicki, and I. Volosyak. 'Autonomous Parameter Adjustment for SSVEP-Based BCIs with a Novel BCI Wizard'. In: *Frontiers in Neuroscience* 9 (Dec. 22, 2015). DOI: 10.3389/fnins.2015.00474.

[87] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz. 'Co-Adaptive Calibration to Improve BCI Efficiency'. In: *Journal of Neural Engineering* 8.2 (Apr. 1, 2011), p. 025009. DOI: 10.1088/1741-2560/8/2/025009.

[88] B. Z. Allison and C. Neuper. 'Could Anyone Use a BCI?' In: *Brain-Computer Interfaces*. Ed. by D. S. Tan and A. Nijholt. London: Springer London, 2010, pp. 35–54. DOI: 10.1007/978-1-84996-272-8_3.

[89] F. Gembler, P. Stawicki, A. Saboor, M. Benda, R. Grichnik, A. Rezeika, and I. Volosyak. 'A Dictionary Driven Mental Typewriter Based on Code-Modulated Visual Evoked Potentials (cVEP)'. In: *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Miyazaki, Japan, Oct. 2018, pp. 619–624. DOI: 10.1109/SMC.2018.00114.

[90] P. Stawicki, F. Gembler, and I. Volosyak. 'Driving a Semiautonomous Mobile Robotic Car Controlled by an SSVEP-Based BCI'. In: *Computational Intelligence and Neuroscience* 2016 (2016), pp. 1–14. DOI: 10.1155/2016/4909685.

[91] K.-K. Shyu, Y.-J. Chiu, P.-L. Lee, J.-M. Liang, and S.-H. Peng. 'Adaptive SSVEP-based BCI system with frequency and pulse duty-cycle stimuli tuning design'. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 21.5 (2013), pp. 697–703. DOI: 10.1109/TNSRE.2013.2265308.

[92] I. Volosyak, T. Malechka, D. Valbuena, and A. Gräser. 'A Novel Calibration Method for SSVEP Based Brain-Computer Interfaces'. In: *Signal Processing Conference, 2010 18th European*. Aalborg, Denmark, 2010, pp. 939–943.

[93] P. Stawicki, F. Gembler, and I. Volosyak. 'Evaluation of Suitable Frequency Differences in SSVEP-Based BCIs'. In: *Symbiotic Interaction: 4th International Workshop, Symbiotic 2015, Berlin, Germany, October 7-8, 2015, Proceedings*. Ed. by B. Blankertz, G. Jacucci, L. Gamberini, A. Spagnolli, and J. Freeman. Cham: Springer International Publishing, 2015, pp. 159–165. DOI: 10.1007/978-3-319-24917-9_17.

[94] N. V. Manyakov, N. Chumerin, A. Robben, A. Combaz, M. van Vliet, and M. M. Van Hulle. 'Sampled Sinusoidal Stimulation Profile and Multichannel Fuzzy Logic Classification for Monitor-Based Phase-Coded SSVEP Brain–Computer Interfacing'. In: *Journal of Neural Engineering* 10.3 (June 1, 2013), p. 036011. DOI: 10.1088/1741-2560/10/3/036011.

[95] A. R. Smith and J. F. Blinn. 'Blue Screen Matting'. In: *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '96*. The 23rd Annual Conference. Not Known: ACM Press, 1996, pp. 259–268. DOI: 10.1145/237170.237263.

[96] X. Chen, Y. Wang, M. Nakanishi, T.-P. Jung, and X. Gao. 'Hybrid Frequency and Phase Coding for a High-Speed SSVEP-Based BCI Speller'. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Chicago, IL, USA, Aug. 2014, pp. 3993–3996. DOI: 10.1109/EMBC.2014.6944499.

[97] M. Nakanishi, Y. Wang, Y.-T. Wang, Y. Mitsukura, and T.-P. Jung. 'Generating Visual Flickers for Eliciting Robust Steady-State Visual Evoked Potentials at Flexible Frequencies Using Monitor Refresh Rate'. In: *PLoS ONE* 9.6 (June 11, 2014), e99235. DOI: 10.1371/journal.pone.0099235.

[98] F. Gembler, P. Stawicki, A. Rezeika, A. Saboor, M. Benda, and I. Volosyak. 'Effects of Monitor Refresh Rates on C-VEP BCIs.' In: *Symbiotic Interaction. Symbiotic 2017. Lecture Notes in Computer Science, Vol 10727.* Ed. by J. Ham, A. Spagnolli, B. Blankertz, L. Gamberini, and G. Jacucci. Cham: Springer, 2018, pp. 53–62. DOI: 10.1007/978-3-319-91593-7_6.

[99] R. C. Reid, J. D. Victor, and R. M. Shapley. 'The Use of M-Sequences in the Analysis of Visual Neurons: Linear Receptive Field Properties'. In: *Visual Neuroscience* 14.06 (Nov. 1997), p. 1015. DOI: 10.1017/S0952523800011743.

[100] A. Mitra. 'On the Properties of Pseudo Noise Sequences with a Simple Proposal of Randomness Test'. In: *International Journal of Electrical and Computer Engineering* 3.3 (2008), pp. 164–169.

[101] Y. Liu, Q. Wei, and Z. Lu. 'A Multi-Target Brain-Computer Interface Based on Code Modulated Visual Evoked Potentials'. In: *PLOS ONE* 13.8 (Aug. 17, 2018). Ed. by B. He, e0202478. DOI: 10.1371/journal.pone.0202478.

[102] B. Wittevrongel, E. Van Wolputte, and M. M. Van Hulle. 'Code-Modulated Visual Evoked Potentials Using Fast Stimulus Presentation and Spatiotemporal Beamformer Decoding'. In: *Scientific Reports* 7.1 (Dec. 2017). DOI: 10.1038/s41598-017-15373-x.

[103] Q. Wei, S. Feng, and Z. Lu. 'Stimulus Specificity of Brain-Computer Interfaces Based on Code Modulation Visual Evoked Potentials'. In: *PLOS ONE* 11.5 (May 31, 2016). Ed. by D. Zhang, e0156416. DOI: 10.1371/journal.pone.0156416.

[104] O. Friman, I. Volosyak, and A. Gräser. 'Multiple Channel Detection of Steady-State Visual Evoked Potentials for Brain-Computer Interfaces'. In: *IEEE Transactions on Biomedical Engineering* 54.4 (Apr. 2007), pp. 742–750. DOI: 10.1109/TBME.2006.889160.

[105] D. Carlson. 'Minimax and Interlacing Theorems for Matrices'. In: *Linear Algebra and its Applications* 54 (Oct. 1983), pp. 153–172. DOI: 10.1016/0024-3795(83)90211-2.

[106] H. Hotelling. 'Relations between Two Sets of Variates'. In: *Biometrika* 28.3-4 (Dec. 1, 1936), pp. 321–377. DOI: 10.1093/biomet/28.3-4.321.

[107] Z. Lin, C. Zhang, W. Wu, and X. Gao. 'Frequency Recognition Based on Canonical Correlation Analysis for SSVEP-Based BCIs'. In: *IEEE Transactions on Biomedical Engineering* 54.6 (2007), pp. 1172–1176. DOI: 10.1109/TBME.2006.889197.

[108] M. Spüler. 'Spatial Filtering of EEG as a Regression Problem'. In: *Proceedings of the 7th Graz Brain-Computer Interface Conference 2017, From Vision to Reality*. DOI: 10.3217/978-3-85125-533-1-84.

[109] Y. Zhang, P. Xu, K. Cheng, and D. Yao. 'Multivariate Synchronization Index for Frequency Recognition of SSVEP-Based Brain–Computer Interface'. In: *Journal of Neuroscience Methods* 221 (Jan. 2014), pp. 32–40. DOI: 10.1016/j.jneumeth.2013.07.018.

[110] V. Mondini, A. L. Mangia, L. Talevi, and A. Cappello. 'Sinc-Windowing and Multiple Correlation Coefficients Improve SSVEP Recognition Based on Canonical Correlation Analysis'. In: *Computational Intelligence and Neuroscience* 2018 (2018), pp. 1–11. DOI: 10.1155/2018/4278782.

[111] M. Nakanishi, Y. Wang, Y.-T. Wang, and T.-P. Jung. 'A Comparison Study of Canonical Correlation Analysis Based Methods for Detecting Steady-State Visual Evoked Potentials'. In: *PLOS ONE* 10.10 (Oct. 19, 2015). Ed. by D. Yao, e0140703. DOI: 10.1371/journal.pone.0140703.

[112] J. Lee Rodgers and W. A. Nicewander. 'Thirteen Ways to Look at the Correlation Coefficient'. In: *The American Statistician* 42.1 (1988), pp. 59–66. DOI: 10.1080/00031305.1988.10475524.

[113] S. Nagel, W. Dreher, W. Rosenstiel, and M. Spüler. 'The Effect of Monitor Raster Latency on VEPs, ERPs and Brain–Computer Interface Performance'. In: *Journal of Neuroscience Methods* 295 (Feb. 2018), pp. 45–50. DOI: 10.1016/j.jneumeth.2017.11.018.

[114] R Zerafa, T Camilleri, O Falzon, and K. P. Camilleri. 'To Train or Not to Train? A Survey on Training of Feature Extraction Methods for SSVEP-Based BCIs'. In: *Journal of Neural Engineering* 15.5 (Oct. 1, 2018), p. 051001. DOI: 10.1088/1741-2552/aaca6e.

[115] I. Volosyak, C. Guger, and A. Gräser. 'Toward BCI Wizard - Best BCI Approach for Each User'. In: *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. 2010 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2010). Buenos Aires, Argentinia, Aug. 2010, pp. 4201–4204. DOI: 10.1109/IEMBS.2010.5627390.

[116] F. Gembler, P. Stawicki, and I. Volosyak. 'Exploring the Possibilities and Limitations of Multitarget SSVEP-Based BCI Applications'. In: *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the the IEEE Engineering in Medicine and Biology Society (EMBC)*. Orlando, FL, USA, 2016, pp. 1488–1491. DOI: 10.1109/EMBC.2016.7590991.

[117] S. Parini, L. Maggi, A. C. Turconi, and G. Andreoni. 'A Robust and Self-Paced BCI System Based on a Four Class SSVEP Paradigm: Algorithms and Protocols for a High-Transfer-Rate Direct Brain Communication'. In: *Computational Intelligence and Neuroscience* 2009 (2009), pp. 1–11. DOI: 10.1155/2009/864564.

[118] X. Chen, Z. Chen, S. Gao, and X. Gao. 'A High-ITR SSVEP-Based BCI Speller'. In: *Brain-Computer Interfaces* 1.3-4 (Oct. 2, 2014), pp. 181–191. DOI: 10.1080/2326263X.2014.944469.

[119] F.-C. Lin, J. K. Zao, K.-C. Tu, Y. Wang, Y.-P. Huang, C.-W. Chuang, H.-Y. Kuo, Y.-Y. Chien, C.-C. Chou, and T.-P. Jung. 'SNR Analysis of High-Frequency Steady-State Visual Evoked Potentials from the Foveal and Extrafoveal Regions of Human Retina'. In: *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. San Diego, CA, USA, 2012, pp. 1810–1814. DOI: 10.1109/EMBC.2012.6346302.

[120] T. Sakurada, T. Kawase, T. Komatsu, and K. Kansaku. 'Use of High-Frequency Visual Stimuli above the Critical Flicker Frequency in a SSVEP-Based BMI'. In: *Clinical Neurophysiology* 126.10 (Oct. 2015), pp. 1972–1978. DOI: 10.1016/j.clinph.2014.12.010.

[121] D.-O. Won, H.-J. Hwang, S. Dähne, K.-R. Müller, and S.-W. Lee. 'Effect of Higher Frequency on the Classification of Steady-State Visual Evoked Potentials'. In: *Journal of Neural Engineering* 13.1 (Feb. 1, 2016), p. 016014. DOI: 10.1088/1741-2560/13/1/016014.

[122] G. G. Molina, D. Ibanez, V. Mihajlović, and D. Chestakov. 'Detection of High Frequency Steady State Visual Evoked Potentials for Brain-Computer Interfaces'. In: *Signal Processing Conference, 2009 17th European*. Glasgow, Scotland, UK, 2009, pp. 646–650.

[123] I. Volosyak, H. Cecotti, and A. Gräser. 'Steady-State Visual Evoked Potential Response - Impact of the Time Segment Length'. In: *Proc. on the 7th International Conference on Biomedical Engineering BioMed2010, Innsbruck, Austria, February 17 –19*. 2010, pp. 288–292. DOI: 10.2316/Journal.216.2010.7.680-0145.

[124] N. S. Dias, P. M. Mendes, and J. H. Correia. 'Subject Age in P300 BCI'. In: *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference On*. Arlington, VA, USA, 2005, pp. 579–582. DOI: 10.1109/cne.2005.1419690.

[125] M. Grosse-Wentrup and B. Schölkopf. 'A Review of Performance Variations in SMR-Based Brain- Computer Interfaces (BCIs)'. In: *Brain-Computer Interface Research*. Springer, 2013, pp. 39–51. DOI: 10.1007/978-3-642-36083-1_5.

[126] H. Macpherson, A. Pipingas, and R. Silberstein. 'A Steady State Visually Evoked Potential Investigation of Memory and Ageing'. In: *Brain and cognition* 69.3 (2009), pp. 571–579. DOI: https://doi.org/10.1016/j.bandc.2008.12.003.

[127] J. Ehlers, D. Valbuena, A. Stiller, and A. Gräser. 'Age-Specific Mechanisms in an SSVEP-Based BCI Scenario: Evidences from Spontaneous Rhythms and Neuronal Oscillators'. In: *Computational Intelligence and Neuroscience* 2012 (2012), pp. 1–9. DOI: 10.1155/2012/967305.

[128] H.-T. Hsu, I.-H. Lee, H.-T. Tsai, H.-C. Chang, K.-K. Shyu, C.-C. Hsu, H.-H. Chang, T.-K. Yeh, C.-Y. Chang, and P.-L. Lee. 'Evaluate the Feasibility of Using Frontal SSVEP to Implement an SSVEP-Based BCI in Young, Elderly and ALS Groups'. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 24.5 (May 2016), pp. 603–615. DOI: 10.1109/TNSRE.2015.2496184.

[129] J. J. S. Norton, J. Mullins, B. E. Alitz, and T. Bretl. 'The Performance of 9–11-Year-Old Children Using an SSVEP-Based BCI for Target Selection'. In: *Journal of Neural Engineering* 15.5 (Oct. 1, 2018), p. 056012. DOI: 10.1088/1741-2552/aacfdd.

[130] F. Gembler, P. Stawicki, and I. Volosyak. 'Towards a User-Friendly BCI for Elderly People'. In: *Proceedings of the 6th International Brain Computer Interface Meeting;* Graz, Austria, 2014. DOI: 10.3217/978-3-85125-378-8-48.

[131] R. Likert. 'A Technique for the Measurement of Attitudes.' In: *Archives of Psychology* 22 140 (1932), pp. 55–55.

[132] J. D. R. Millán. 'Combining Brain-Computer Interfaces and Assistive Technologies: State-of-the-Art and Challenges'. In: *Frontiers in Neuroscience* 1 (2010). DOI: 10.3389/fnins.2010.00161.

[133] E. Holz, L. Botrel, and A. Kübler. 'Bridging Gaps Long-Term Independent BCI Home-Use by a Locked-in End-User'. In: TOBI Workshop IV. Sion, Switzerland, Jan. 1, 2013.

[134] T. Kaufmann, S. M. Schulz, A. Köblitz, G. Renner, C. Wessig, and A. Kübler. 'Face Stimuli Effectively Prevent Brain–Computer Interface Inefficiency in Patients with Neurodegenerative Disease'. In: *Clinical Neurophysiology* 124.5 (May 2013), pp. 893–900. DOI: 10.1016/j.clinph.2012.11.006.

[135] A. Kübler, E. M. Holz, E. W. Sellers, and T. M. Vaughan. 'Toward Independent Home Use of Brain-Computer Interfaces: A Decision Algorithm for Selection of Potential End-Users'. In: *Archives of Physical Medicine and Rehabilitation* 96.3 (Mar. 2015), S27–S32. DOI: 10.1016/j.apmr.2014.03.036.

[136] A. Riccio, L. Simione, F. Schettini, A. Pizzimenti, M. Inghilleri, M. O. Belardinelli, D. Mattia, and F. Cincotti. 'Attention and P300-Based BCI Performance in People with Amyotrophic Lateral Sclerosis'. In: *Frontiers in Human Neuroscience* 7 (2013). DOI: 10.3389/fnhum.2013.00732.

[137] E. W. Sellers, T. M. Vaughan, and J. R. Wolpaw. 'A Brain-Computer Interface for Long-Term Independent Home Use'. In: *Amyotrophic Lateral Sclerosis* 11.5 (Oct. 2010), pp. 449–455. DOI: 10.3109/17482961003777470.

[138] Y. Punsawad and Y. Wongsawat. 'Minimal-Assisted SSVEP-Based Brain-Computer Interface Device'. In: *Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference*. Hollywood, CA, USA, Dec. 2012, pp. 1–4.

[139] P.-J. Kindermans and B. Schrauwen. 'Dynamic Stopping in a Calibration-Less P300 Speller'. In: *International BCI Meeting Brain-Computer Interface 2013. Proceedings of the Fifth International Brain-Computer Interface Meeting: Defining the Future. June 3-7 2013 Asilomar Conference Center, Pacific Grove, California, USA* (2013). DOI: 10.3217/978-3-85125-260-6-75.

[140] P.-J. Kindermans, M. Schreuder, B. Schrauwen, K.-R. Müller, and M. Tangermann. 'True Zero-Training Brain-Computer Interfacing – An Online Study'. In: *PLoS ONE* 9.7 (July 28, 2014), e102504. DOI: 10.1371/journal.pone.0102504.

[141] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller. 'How Many People Are Able to Operate an EEG-Based Brain-Computer Interface (BCI)?' In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 11.2 (June 2003), pp. 145–147. DOI: 10.1109/TNSRE.2003.814481.

[142] C. Guger, B. Z. Allison, B. Großwindhager, R. Prückl, C. Hintermüller, C. Kapeller, M. Bruckner, G. Krausz, and G. Edlinger. 'How Many People Could Use an SSVEP BCI?' In: *Frontiers in Neuroscience* 6 (2012). DOI: 10.3389/fnins.2012.00169.

[143] T. M. S. Mukesh, V Jaganathan, and M. R. Reddy. 'A Novel Multiple Frequency Stimulation Method for Steady State VEP Based Brain Computer Interfaces'. In: *Physiological Measurement* 27.1 (Jan. 1, 2006), pp. 61–71. DOI: 10.1088/0967-3334/27/1/006.

[144] M. A. Pastor, J. Artieda, J. Arbizu, M. Valencia, and J. C. Masdeu. 'Human Cerebral Activation during Steady-State Visual-Evoked Responses'. In: *The Journal of Neuroscience* 23.37 (Dec. 17, 2003), pp. 11621–11627. DOI: 10.1523/JNEUROSCI.23-37-11621.2003.

[145] Y.-T. Wang, Y. Wang, and T.-P. Jung. 'A Cell-Phone-Based Brain–Computer Interface for Communication in Daily Life'. In: *Journal of Neural Engineering* 8.2 (Apr. 1, 2011), p. 025018. DOI: 10.1088/1741-2560/8/2/025018.

[146] S. N. Carvalho, T. B. S. Costa, L. F. S. Uribe, D. C. Soriano, S. R. M. Almeida, L. L. Min, G. Castellano, and R. Attux. 'Effect of the Combination of Different Numbers of Flickering Frequencies in an SSVEP-BCI for Healthy Volunteers and Stroke Patients'. In: *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*. 2015 7th International IEEE/EMBS Conference on Neural Engineering (NER). Montpellier, France, Apr. 2015, pp. 78–81. DOI: 10.1109/NER.2015.7146564.

[147] R. Fazel-Rezai. 'Human Error in P300 Speller Paradigm for Brain-Computer Interface'. In: *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. Lyon, France, 2007, pp. 2516–2519. DOI: 10.1109/iembs.2007.4352840.

[148] F. Gembler, P. Stawicki, A. Rezeika, and I. Volosyak. 'A Comparison of cVEP-Based BCI-Performance Between Different Age Groups'. In: *Advances in Computational Intelligence*. Ed. by I. Rojas, G. Joya, and A. Catala. Vol. 11506. Cham: Springer International Publishing, 2019, pp. 394–405. DOI: 10.1007/978-3-030-20521-8_33.

[149] F. Gembler, P. Stawicki, A. Saboor, and I. Volosyak. 'Dynamic Time Window Mechanism for Time Synchronous VEP-Based BCIs—Performance Evaluation with a Dictionary-Supported BCI Speller Employing SSVEP and c-VEP'. In: *PLOS ONE* 14.6 (June 13, 2019). Ed. by Z. Wang, e0218177. DOI: 10.1371/journal.pone.0218177.

[150] F. Gembler, M. Benda, A. Saboor, and I. Volosyak. 'A Multi-Target c-VEP-Based BCI Speller Utilizing n-Gram Word Prediction and Filter Bank Classification'. In: *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). Oct. 2019, pp. 2719–2724. DOI: 10.1109/SMC.2019.8914235.

[151] T. Eckart and U. Quasthoff. 'Statistical Corpus and Language Comparison on Comparable Corpora'. In: *Building and Using Comparable Corpora*. Ed. by S. Sharoff, R. Rapp, P. Zweigenbaum, and P. Fung. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 151–165. DOI: 10.1007/978-3-642-20128-8_8.

[152] F. Gembler, P. Stawicki, and I. Volosyak. 'How Many Electrodes Are Needed for Multi-Target SSVEP-BCI Control: Exploring the Minimum Number of Signal Electrodes for CCA and MEC'. In: *Proceedings of the 7th Graz Brain-Computer Interface Conference 2017, From Vision to Reality*. 2017. DOI: 10.3217/978-3-85125-533-1-29.

[153] M. Middendorf, G. McMillan, G. Calhoun, and K. S. Jones. 'Brain-Computer Interfaces Based on the Steady-State Visual-Evoked Response'. In: *IEEE transactions on rehabilitation engineering* 8.2 (2000), pp. 211–214. DOI: 10.1109/86.847819.

[154] M. Benda, P. Stawicki, F. Gembler, R. Grichnik, A. Rezeika, A. Saboor, and I. Volosyak. 'Different Feedback Methods For An SSVEP-Based BCI'. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Honolulu, HI, USA, July 2018, pp. 1939–1943. DOI: 10.1109/EMBC.2018.8512622.

[155] C. Jia, X. Gao, B. Hong, and S. Gao. 'Frequency and Phase Mixed Coding in SSVEP-Based Brain–Computer Interface'. In: *IEEE Transactions on Biomedical Engineering* 58.1 (Jan. 2011), pp. 200–206. DOI: 10.1109/TBME.2010.2068571.

[156] F. Gustafsson. 'Determining the Initial States in Forward-Backward Filtering'. In: *IEEE Transactions on Signal Processing* 44.4 (Apr. 1996), pp. 988–992. DOI: 10.1109/78.492552.

[157] Y.-Y. Chien, F.-C. Lin, J. K. Zao, C.-C. Chou, Y.-P. Huang, H.-Y. Kuo, Y. Wang, T.-P. Jung, and H.-P. D. Shieh. 'Polychromatic SSVEP Stimuli with Subtle Flickering Adapted to Brain-Display Interactions'. In: *Journal of Neural Engineering* 14.1 (Feb. 1, 2017), p. 016018. DOI: 10.1088/1741-2552/aa550d.

# Notation

| | |
|---|---|
| ADC | Analog-to-digital converter |
| ALS | Amyotrophic lateral sclerosis |
| BCI | Brain-computer interface |
| c-VEP | Code-modulated visual evoked potential |
| CCA | Canonical correlation analysis |
| CLM | Correct letters per minute |
| CNS | Central nervous system |
| CRT | Cathode-ray tube |
| ECoG | Electrocorticography |
| EEG | Electroencephalography |
| ERD | Event-related desynchronization |
| ERP | Event-related potential |
| ERS | Event-related synchronization |
| f-VEP | Frequency-modulated visual evoked potential |
| FFT | Fast Fourier transform |
| fMRI | Functional magnetic resonance imaging |
| FN | False negative |
| FP | False positive |
| GUI | Graphical user interface |
| ITR | Information transfer rate |
| LCD | Liquid-crystal display |
| LFSR | Linear-feedback shift register |
| LIS | Locked-in syndrome |
| m-VEP | Motion-onset visual evoked potential |

MEC Minimum energy combination

MEG Magnetoencephalography

MI Motor imagery

MND Motor neuron disease

NIRS Near infrared spectroscopy

OCM Output characters per minute

PCA principal component analysis

PCM Pearson's correlation method

PSD Power spectrum density

SMI Similarity index

SMR Sensorimotor rhythms

SNR Signal-to-noise ratio

SQL Structured query language

SSMVEP Steady-state motion visual evoked potential

SSVEP Steady-state visual evoked potential

TN True negative

TP True positive

VEP Visual evoked potential

# Declaration

I declare that I have completed this work solely and only with the help of the references mentioned.

*Goch-Pfalzdorf, November 3, 2020*

Felix Gembler