University of Osnabrück

Analyzing event-related potentials in 8-channel EEG data using machine learning methods

Ann-Kathrin Schalkamp
963930

A thesis presented for the degree of
Bachelor of Science

First Supervisor: Olivera Stojanovic, M.Sc.
Second Supervisor: Prof. Dr. Gordon Pipa

26.09.2018

Abstract

Over the past few years Brain Computer Interface (BCI) research programs have arisen searching for a new channel of non-muscular communication with the external world. One possible channel can be provided via electroencephalography (EEG) systems, which probe the electric potential on the scalp with electrodes. EEG systems are becoming cheaper, portable versions are developed, and such systems are often used as the recording system for BCIs. In this study a recently developed portable high-tech sleep mask, the *Traumschreiber*, was used to record ERPs. The placement of the electrodes was performed in a different way compared to standard EEG systems. After re-referencing only one recording electrode but several references produced data for eight channels. Different combinations of the data were used for training logistic regression models (single channel, best three channels to increase dimensionality, best three channels to increase training set). Those were evaluated based on their predictive performance by using the area under the curve (AUC) as a metric. As a further analysis the data of subjects was combined and compared to the results of the subject specific analysis and the traditional averaging method.
In this study the usage of the single best channel resulted in the best prediction performance (0.74 for SSVEP, 0.7 for MP, 0.76 for N240 targets and 0.76 for N240 distractors). Combining data of several channels, that use the same recording electrode to increase the amount of data produced good results regardless of the task's complexity (0.72 for SSVEP, 0.66 for MP, 0.76 for N240 targets and 0.76 for N240 distractors). For the data of low-level tasks the combination of several subjects proved to be an option to increase the training data as well (0.73 for SSVEP and 0.66 for MP).

List of Figures

List of Tables

List of Abbreviations

AUC area under the curve

BCI Brain Computer Interface

EEG electroencephalography

ECG electrocardiography

EOG electrooculography

ERP event-related potential

fMRI functional magnetic resonance imaging

HCI Human Computer Interface

i.i.d. independent and identically distributed

IRLS iteratively reweighted least squares

LSE least squares error

MLE Maximum Likelihood Estimation

MP motor potential

N240 negativity after 240 ms

RAP reafferent potential

ROC receiver-operator characteristic

SSVEP steady-state evoked potential

vDM visual decision making

Contents

1 Introduction

Electroencephalography is one of the most commonly used techniques in noninvasive brain research (Melnik et al., 2017). Over the last years there has been movement in related research areas due to the improvement of electrical components. Those components are getting more powerful, cheaper, and smaller at the same time (Appel, 2017). This allows for the construction of mobile EEG systems which can address new research questions. Mobile EEG systems allow to examine cognition under natural contexts and cognitive processes involved in actions and sensorimotor tasks (Melnik et al., 2017).
Most mobile EEG projects are open source and their knowledge is provided to everyone who is interested, e.g. OpenEEG (OpenEEGProject, 2018). This allows for more exchange of information, faster improvements and lower costs.
As a consequence there are plenty of new low-cost EEG systems that claim to be as powerful as standard EEG systems while allowing free movement at the same time. To test these claims, quantitive studies have been carried out comparing EEG systems (Oliveira et al., 2016; Melnik et al., 2017). Those comparative studies show that the choice of the EEG system influences the recorded data and that systems are a significant source of variance (Melnik et al., 2017).
In the field of BCIs EEG systems are commonly used. Performing single-trial analysis on EEG data, meaning an online classification based on the data of a single event-related potential (ERP), allows for the translation of the brain's activity into commands for the external world. Studies only using a few channels can perform single-trial analysis of EEG data (Gao et al., 2003; Mandel et al., 2009). In order to classify the EEG data machine learning techniques are applied. They need to be able to deal with data which has a poor signal-to-noise ratio and mislabels. Various different techniques proved to be successful (Lotte et al., 2007).

Because of the rise of mobile EEG systems and the advances in the field of BCIs it may be time to question some standards established over years of EEG research. Mobile EEG systems and other EEG systems used for BCIs often use a smaller number of electrodes compared to the standard EEG systems. Some do not use the standard placement of electrodes. Others use dry electrodes. Hence set standards like the usage of as much electrodes as possible, a commonly used placement of electrodes and the reference or the filtering of data may not apply for BCI applications. Recent BCI and mobile EEG studies showed that EEG analysis can be performed without compliying with the standards. Rethinking those standards might even lead to new ideas to deal with challenges of BCI research.

1.1 Brain Computer Interfaces

HCIs are already well established and used daily but an increasing need for direct communication with external devices led to the rise of a new research field, BCIs (Vallabhaneni, Wang, & He, 2005). Direct communication is needed if the brain's normal neuromuscular output channel is impaired.
BCIs try to establish a direct pathway between a brain and an external device. The direct communication via neuronal activity renders the communication via peripheral pathways unnecessary making external devices more usable for motion disabled people (Gao et al.,

2003). BCIs may serve as a replacement of the brain's normal neuromuscular output channels (Wolpaw et al., 2002). This additional flow of information can therefore help to repair human functions or add other features that can improve human cognition. BCI research focuses mainly on medical applications but recently a variety of low cost BCIs were also built into toys and gaming devices.

The new form of communication does not take place by determining a person's intentions or needs, but a new channel is built which can be voluntarily controlled by the user. BCIs can either be noninvasive and record brain signals from outside the brain or invasive like intracortical BCIs. The best studied noninvasive BCIs are EEG-based.

Today's EEG-based BCIs allow a communication via electrophysiological signals (Wolpaw et al., 2002). The user voluntarily generates a certain neuronal potentials that is translated in real-time into a command for the computer. In order to allow the identification of those signals, an initial adaptation of user and interface is needed as well as an ongoing adjustment (Wolpaw et al., 2002).

A BCI consists of three elements (Wolpaw et al., 2002). Firstly, the user's input is recorded. EEG-based BCIs record brain activity from the scalp. EEG will be discussed further in the next chapter (1.2). Secondly, the recorded signal is processed and relevant features are extracted and translated into commands. For EEG-based interfaces these features can be in the time or frequency domain. This step is mostly carried out using machine learning approaches. More on that will follow in the chapter about EEG analysis (1.4). Thirdly, a protocol that determines onset, offset and timing of the output is needed. This takes care of a secure control and allows an easy handling.

Various EEG-based BCIs are already developed and tested. BCI competitions hosted by the Berlin Brain-Computer Interface team provide EEG data with the purpose of comparing data analysis methods. Different challenges for BCI systems are addressed like having small training sets or data that is affected by eye movement or dealing with the problem of session-to-session transfer (Berlin Brain - Computer Interface, 2018). Some publications related to these data sets were able to overcome these challenges (Wang et al., 2004; Lemm et al., 2004; Kaper et al., 2004) but the search for the best analysis techniques is ongoing.

1.2 Electroencephalography

Currently, EEG is widely used in noninvasive brain research as a technique to measure brain activity that is present on the scalp. Research questions related to EEG aim at finding the correlates of "perceptual, cognitive, and motor activity associated with processing of information" (Melnik et al., 2017).

Measuring brain activity at the scalp is only possible when many neurons form a population by being aligned in parallel orientation and firing synchronously (Cohen, 2014, p. 51). This synchronous activity of at least hundreds of neurons allows EEG systems to detect rhythmic changes of electrical activity at the scalp (Cohen, 2014, p. 51). The activity that is present on the scalp is mainly the summation of excitatory and inhibitory postsynaptic potentials at the dendrites (Cohen, 2014, p. 51). The generated electric field is the result of ion flow in and out of each cell such that the membrane potential changes. The sum of the electric fields is recorded by the electrodes placed on the scalp.

Since the electrodes are placed outside the brain there is some air between the electrodes

and the scalp. In order to overcome the issue that air is electrically non-conducting, the gap between scalp and electrodes is closed by an electroconductive gel or salt-water (Cohen, 2014, p. 51). Electroconductive gel or salt-water-soaked sponges are used to form this bridge between each electrode and the scalp.
In the last years EEG became the most used technique for noninvasive BCIs. This is due to the many advantages of EEG systems.

- As Cohen states, a major benefit of EEG is its *high temporal precision* and accuracy. The downside is that it has a rather low spatial resolution compared to functional magnetic resonance imaging (fMRI) and other brain imaging techniques (Cohen, 2014, p. 26). A high temporal resolution allows an online evaluation of the potentials which can be used as an input to BCIs. Most cognitive processes are fast and occur within tens of hundreds of milliseconds. Oftentimes these processes form a sequence of up to a few seconds. Temporal precision enables the system to capture these sequences in the right order. Therefore, Cohen concludes that high temporal resolution techniques are suitable to record signals that are dynamic and temporally sequenced like cognitive events (Cohen, 2014, p. 15).

- EEG also *directly measures brain activity* and does not use other biophysical phenomena which are related to brain activity. E.g., fMRI uses the hemodynamic response in order to infer brain activity. The recorded voltage fluctuations, that are measured by EEG systems, directly reflect the brain activity (Cohen, 2014, p. 15).

- Another reason to use EEG is the *multidimensionality* of its recorded data. According to Cohen, it represents at least four dimensions: "time, space, frequency, and power (the strength of frequency-band-specific activity) and phase (the timing of the activity)" (Cohen, 2014, p. 16). This variety of the data allows for more advanced analysis techniques.

1.2.1 Mobile EEG

Low cost, mobile EEG systems are highly requested for various research fields. That is because of some shortcomings of standard EEG systems and the better usability of portable devices for BCIs.

- EEG studies can *cost* quite a lot even though the same system can be used several times. Cohen estimated the costs for an EEG study with 250$ per subject. But with even more equipment like eye trackers or powerful computers the costs can easily rise to 375$ (Cohen, 2014, p. 29).

- The *preparation* of the subjects is quite time-consuming. At least one researcher has to be present to apply the gel. Caps with up to 128 electrodes are used in present studies and thus the preparation time can expand to more than half an hour.

- For most studies the participants are told to move as least as possible and even avoid eye blinks. This is due to the fact that otherwise artifacts like muscle movement are recorded. But many cognitive processes involve *movement* of the subject or interaction with the environment. Research questions that address those processes cannot be examined with standard EEG systems since those do not allow free movement.

Due to these shortcomings mobile EEG devices are currently developed and often used for BCIs. Those mostly open source projects try to build low-cost, easily applicable systems that allow for research which involves movement. Oftentimes they aim at systems that use less powerful devices for recording and Bluetooth for transmission to support free movement. Less powerful devices might not be capable of setting highly accurate markers. Therefore Krigolson et al. proposed an alternative to continuous EEG data. They performed an EEG study without setting markers and using only a single computer. They did so by recording for only a short amount of time after each stimulus presentation (Krigolson et al., 2017).
Those new approaches seem to be really helpful but unfortunately there is no consent on how to test whether those new EEG devices are capable of producing the same results as the standard ones. Melnik et al. proposed a benchmark for new EEG systems (Melnik et al., 2017). This is highly relevant since the used EEG system contributes to the variance by the same order of magnitude as subjects do. Mobile EEG systems should be able to produce comparable results to standard EEG systems. According to Melnik et al., mobile EEG systems only do so to a varying degree. But other studies claim that less electrodes, non-standardized placement of electrodes and a moving subject still provide good results (Krigolson et al., 2017).

1.3 Event-Related Potentials

Melnik et al. propose a benchmark for new systems by means of ERPs. Those potentials are averaged EEG responses aligned to a stimulus. ERPs reflect how the processing of a particular stimulus affects the general activity of the brain (Anderson, 2015, p. 21). In order to eliminate the effects that are not in direct relation to the stimulus processing the average over many trials and subjects is calculated. The remaining change of potential reflects the activity resulting from the stimulus processing (Anderson, 2015, p. 21). For example the N240 component is a negativity that can be observed about 240 ms after the stimulus onset. ERPs in general have a very good temporal resolution but the interference of the location of the potential is rather difficult.
Mobile EEG systems are not designed to investigate subtle ERPs only visible in lab conditions. Instead they allow for the investigation of cognitive processes involved in movement. Those show in everyday life. To take this into account, prominent ERPs should be used to form a benchmark (Melnik et al., 2017).
Following the proposal of Melnik et al. the study conducted for this thesis was designed. The experiments were executed by Merle Reimann and Ann-Kathrin Schalkamp. The examined ERPs were motor potentials, steady-state visual evoked potentials, auditory potentials, visual mismatch negativity, P300 and visual decision making/N240. In this thesis steady-state visual evoked potentials (SSVEPs), motor potentials (MPs) and visual decision making (vDM/N240) will be analysed.

1.3.1 SSVEP

Steady-state visual evoked potentials are a kind of flicker effect. It describes the phenomenon of the brain activity becoming entrained to a rhythmic factor. This shows in the ERP although it takes several milliseconds for the effect to stabilize (Cohen, 2014, p. 101). Factors that influence the flicker effect are the frequency at which the stimulus is presented and also the size, contrast and luminance of the stimulus will alter the effect. Lower frequencies elicit a stronger flicker effect but also a stimulus frequency of up to 100 Hz can evoke a steady-state evoked potential (SSVEP) (Herrmann, 2001). SSVEPs are typically recorded at medial occipital regions near the Oz electrode (Norcia et al., 2015). That iy why most researchers concluded that the SSVEP originates in the primary visual cortex and reflects low-level vision (Di Russo et al., 2007). However, SSVEPs can also be used to study higher level visual processes because it varies according to frequency and the stimulus type (Norcia et al., 2015).

1.3.2 Motor Potential

A voluntary movement action normally involves a readiness potential, the motor potential (MP) itself and a reafferent potential (RAP). The MP shows during the movement and the RAP follows such a voluntary motor action. These potentials are assumed to be maximal above the motor cortex and biased towards the contralateral side (Cacioppo et al., 2007, p. 95). Hence, electrodes placed above the precentral sulcus like AFz, Cz, FC3 and FC4 should record the signal the best.

1.3.3 Visual Decision Making: N240

The N240 component can be elicited by paradigms where the participant has to actively attend the stimuli and react only to the deviant one (Patel & Azzam, 2005). The component belongs to the N2 family where a negativity between 180 to 325 ms results from the deviation from the prevalent stimulus (Patel & Azzam, 2005). In the case of the N240 component the negativity shows about 240 ms after the stimulus onset. The N2 family comprises three different types of components: N2a, N2b and N2c (Sur & Sinha, 2009). The N240 belongs to the group of N2b and can be elicited by oddball detection. The component can be recorded around the left parietal cortex near the PO3 electrode (Melnik et al., 2017). But as visual processing is involved, the signal can also be recorded above the visual cortex located in the occipital lobe. Oftentimes the difference between the recorded potentials for target and distractor is calculated.

1.4 Analysis of ERPs

A major challenge for the analysis of ERPs is the low signal-to-noise ratio of the data. The traditional way to overcome this challenge is to average over many trials. The waveforms elicited by each shown stimulus will be averaged. This helps to remove the noise components because they appear randomly. But BCIs aim at single-trial detection of ERPs. For this task more powerful signal processing and classification approaches are needed (Huang et al., 2006).

The classification algorithm needs to be able to deal with high dimensionality, noise and outliers, and small training sets (Lotte et al., 2007). The high dimensionality arises from the usage of several electrodes and a high sampling rate. Noise and other interfering factors lead to the problem that in this high dimensional space the relevant information is concentrated in a low dimensional subspace. That is why a feature extraction has to be performed (Lemm et al., 2011). Fewer electrodes and a lower sampling rate can help to simplify this preprocessing step. Since, according to Lotte et al., the application of a BCI should be fast and easy, only a few data points are available for training (Lotte et al., 2007). Thus, the classifier should be able to learn the underlying structure of the signal fast and efficiently.

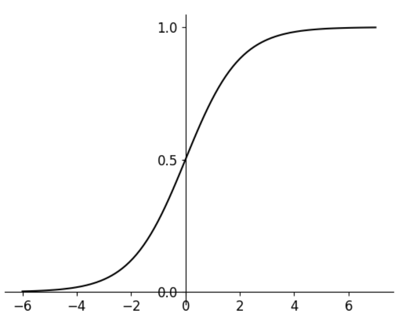
1.4.1 Logistic Regression Classifier

Logistic regression solves binary classification fast and effectively. In contrast to other regression algorithms it does not predict a continuous outcome but a binary one instead. Hence, it is used to predict a categorical outcome from independent variables. For EEG data logistic regression uses the assumption that the EEG signal is a "linear combination of distributed source activity and zero-mean white Gaussian noise measurement" (Huang et al., 2006). Therefore, the strategy to detect ERPs is to the find optimal linear projections of the sensor measurements (Huang et al., 2006). The goal is to maximize the discrimination abilities.

In order to produce a binary output, the linear regression equation (1) is modified by applying a sigmoid function (3). The logistic function or sigmoid function is used to transform the predicted output into the interval between 0 and 1. These output values can either be used as probabilities of belonging to a class, or a class membership can be determined using a threshold. The logistic function is S-shaped as can be seen in figure 1.1.

$$\hat{y} = \mathbf{w}^\top \mathbf{x} + b \tag{1}$$

$$\hat{y} = \sigma(\mathbf{w}^\top \mathbf{x} + b) \tag{2}$$



$$\sigma(z) = \frac{1}{1 + e^{-z}} = \frac{z}{1 + e^{z}} \tag{3}$$

Figure 1.1: Logistic Function

A discriminative classifier like logistic regression tries to directly estimate parameters of the posterior distribution, $P(Y|X)$ which is the probability of the class given the independent

parameters. This posterior distribution is assumed to be given by the equation (4).
When the data is projected to one coordinate, it should be distributed according to a logistic function (Parra et al., 2005) like the one in figure 1.1. The linear discriminant function (1) consists of y the real class label, $\mathbf{w}$ the projection vector, $\mathbf{x}$ the data and b the bias. The two classes get separated by a hyperplane orthogonal to $\mathbf{w}$ and shifted by b. With the named assumptions and the linear discriminant function, one can calculate the likelihood of a sample x_i being in the class of positive examples:

$$P(y_i = 0|x_i) = \frac{e^{\hat{y}_i}}{1 + e^{\hat{y}_i}} = \frac{e^{\mathbf{w}^\top x_i + b}}{1 + e^{\mathbf{w}^\top x_i + b}} \tag{4}$$

Maximum Likelihood Estimation (MLE) is used to approximate the regression parameters in order to maximize the correctly classified new examples. Thus, to find suitable parameters, the likelihood of the data is maximized with respect to the model parameters. Those regression parameters, $\mathbf{w}$ and b are chosen to be the ones that give the highest probabilities for the actual class.

$$\max_\theta P_\theta((y_1, y_2, ..., y_t)|(x_1, x_2, ..., x_t)) \tag{5}$$

The examples are assumed to be independent of each other, thus the equation can be rewritten into

$$\max_\theta P_\theta(y_1|x_1) P_\theta(y_2|x_2) ... P_\theta(y_t|x_t). \tag{6}$$

To make the function easier to work with, one can transform it into the log likelihood

$$\max_\theta \log P_\theta(y_1|x_1) + \log P_\theta(y_2|x_2) + ... + \log P_\theta(y_t|x_t). \tag{7}$$

To obtain a proper loss function, the sign is flipped and the negative log-likelihood function is used.

$$\min_\theta(-\sum_{i=1}^{t} \log P(y_i|x_i)) \tag{8}$$

If $\hat{y}_i$ is the probability of $\mathbf{x_i}$ belonging to the positive class "0", the probability of it belonging to the negative class "1" is $1 - \hat{y}_i$.

$$P_\theta(y_i|x_i) = \begin{cases} \hat{y}_i, & \text{if} y_i = 0 \\ 1 - \hat{y}_i, & \text{if} y_i = 1 \end{cases} \tag{9}$$

Thus the resulting learning objective is

$$l(y, \hat{y}) = -\sum_{i=1}^{t} y_i \log 1 - \hat{y}_i + (1 - y_i) \log \hat{y}_i. \tag{10}$$

As y_i is either 0 or 1 one of the terms will vanish. On the one hand, if it is 0 the term $\log \hat{y}_i$ remains, meaning that the probability of x_i belonging to the positive class should be maximized. If y_i equals 1 on the other hand, $\log 1 - \hat{y}_i$ remains being the probability of x_i belonging to the negative class.
A fast and efficient way to solve this optimization problem is to use the iteratively reweighted least squares (IRLS) algorithm (McCullagh & Nelder, 1989).

The reason why logistic regression is well suited for single-trial analysis classification is that it does not require many computational resources, the input features do not need to be scaled and it can be easily regularized. Logistic regression works best if unimportant features and correlated ones are removed. Therefore epochs that comprise the informative features of the ERP should be extracted. Regularization can be used to deal with its sensitivity to overfitting. Logistic regression is sensitive to misclassified examples which is problematic as those might be included in the labelled EEG data. It has been shown that single-trial analysis of EEG data works well using logistic regression and that it can even outperform other classifiers (Tomioka et al., 2007; Bashashati et al., 2015).

1.4.2 Regularization

A lot of machine learning methods suffer from overfitting. Overfitting is defined as achieving the best results for the training data but performing worse on the test data. This occurs when the model picks up noise or peculiarities. The model then fails to generalize.
In order to improve generalization regularization can be used. Regularization is a form of penalization of complexity. Therefore, it is useful if the data has high variance. It introduces a bias to the algorithm that reduces the generalization error. L1 or L2 regularization can be applied. L2 regularization is more computationally efficient and thus better suited for a fast online classification.
L2 regularization adds a regularization term to the cost function of the logistic regression (see equation (10)) as defined in equation (11). This introduces the least squares error (LSE). Each of the n coefficients is adjusted in the same way. This additional constraint increases as the coefficients of the model increase. The new hyperparameter λ controls the regularization strength. Higher values for λ lead to smaller coefficients, but too high values can result in underfitting. Thus, the choice of this hyperparameter has to be thought through.

$$L2 : \frac{\lambda}{2} \sum_{j=1}^{n} w_j^2 \tag{11}$$

1.5 Aim of the Study

This study is performed to investigate whether EEG data acquired with the *Traumschreiber* is usable for traditional and single-trial analysis. Moreover, the importance of the placement of electrodes will be examined. To do so, the benchmark consisting of six paradigms proposed by Melnik et al. was used as an orientation (Melnik et al., 2017).
The ERP components will be evaluated by using the standard averaging method as well as logistic regression as a single-trial analysis method. Different channels and different combinations of those as input vectors for the logistic regression models will be compared by their accuracy on the test data. By doing so this study aims at showing that standards of EEG research might not be useful for BCIs.

2 Materials and Methods

The *Traumschreiber* developed by Johannes Leugering and Kristoffer Appel from the University of Osnabrück under the supervision of Prof. Dr. Gordon Pipa served as a low cost mobile EEG system. The *Traumschreiber* was designed to measure and influence sleep. Some studies already successfully used this device for polysomnography (Appel, 2017; Mandt, 2017). Instead of connecting electrodes for electrocardiography (ECG) and electrooculography (EOG), ten ring electrodes were connected to the *Traumschreiber* and attached to an EEG cap, the Easy-Cap (Brain Products, 2018). A connection to a computer, Lenovo IdeaPad U330 Touch was established via Bluetooth. This computer using Ubuntu 18.04.1 LTS as its operating system served as the stimulus generator, received the data from the *Traumschreiber* and set markers. The code for the *Traumschreiber*, the liveplotting function (see 2.4) and the framework for the stimulus generation and data storage were provided by Johannes Leugering using python 3.6 and its pygame library.

2.1 Participants

Nine healthy participants were recruited per e-mail announcement or orally. By mistake a different placement of electrodes was performed for subject 3. This is why that subject was left out of the evaluation. Therefore, the data of eight participants (1 male, 7 females, mean age = 20.63 years, range: 19 - 21 years) will be analysed in this study. Students of the Univerity of Osnabrück could receive two test-person hours which some of them need for their examination. All of them were native German speakers. Exclusion criteria were being bald, having dreadlocks or suffering from neurological disorders like epilepsy. Written consent was obtained prior to the study and the participants were free to quit the experiment at any point without giving reasons. An exemplar of the consent form can be found in the appendix A. The experiment was approved by the local ethics committee of the University of Osnabrück.

2.2 Traumschreiber

The *Traumschreiber* is still in a developmental stage and several versions of it exist for different purposes. In this study a battery-driven system with ten connections for electrodes was used (see figure 2.1). This setup allows for the measurement of eight EEG channels, as the *Traumschreiber* uses bipolar electrodes. Therefore, the difference between a pair of electrodes is measured. Such a setup was chosen for the *Traumschreiber*, as its original purpose is to measure EOG and ECG where bipolar electrodes are commonly used.

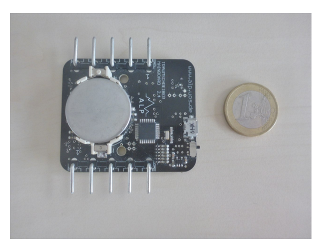


Figure 2.1: Used version of *Traumschreiber*

The recorded data was transmitted to the computer via Bluetooth which causes a buffering. Python scripts provided by Johannes Leugering stored the data. The stimulus generating code also set markers when a stimulus was shown such that the continuous data can later on be epoched into meaningful units. A sampling rate of 220 Hz and a gain of one were used for recording. The

Traumschreiber is able to record 244 data points every second but on average 220 points are transferred (Appel, 2017). The gain can also be set to up to 64.
As there is no impedance measurement available for the *Traumschreiber*, instead a liveplot of the data was used to get an impression of the data. Thereby, bad electrodes can be located and the connection can be improved when needed (see figure 2.4). In standard EEG studies the impedances should be kept below 5 or 10 kOhm as high electrode impedances can increase noise and thus reduce statistical power. However, some studies showed that a cool and dry environment can reduce these effects (Kappenman & Luck, 2010). Therefore, such an environment was provided as a prior adjustment since the impedances were unknown and could be high.

2.3 Design

The chosen paradigms are reproductions of the ones used by Melnik et al. in their study. Therefore, the same stimuli were presented under the same conditions but with minor adjustments which will be specified below. The total amount of trials was reduced as more subjects participated in this study than the one of Melnik et al. and the recorded data should suffice for analysis.

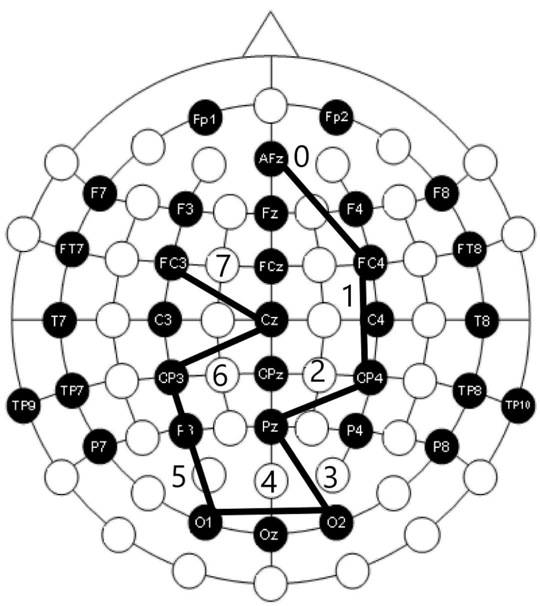


Figure 2.2: EasyCap with used electrode placement. Adapted from (FieldTrip, 2018).

The electrodes were placed over regions where the literature postulates the investigated ERPs are maximal. Therefore Fz, Pz, Cz, AFz, FC4, FC3, CP3, CP4, O1 and O2 were chosen as well as TP10 as the ground electrode. Since the *Traumschreiber* records pairwise differences in electrical activity, the pairwise connections were chosen in a way to maximize spreading in parietal and occipital regions where the ERPs are assumed to be most prominent. The pairwise couplings can be seen in figure 2.2.

The paradigms were presented in a random order in order to reduce the effects of decreasing concentration and increasing tiredness. The overall experiment consisted of thirty minutes of preparation and one hour of recording with a break halfway through. The participants were also free to choose when to start each experiment so they could take individual breaks throughout the whole experiment.
During the experiment the participants were seated in a darkened chamber in front of a monitor, an audio box and a keyboard. The *Traumschreiber* was placed in a small bag which was attached to the back of the participants' chair. The computer, which ran the code, was placed outside the experiment chamber and was connected to the *Traumschreiber* via Bluetooth and to the other materials via HDMI.

2.3.1 SSVEP

The same stimuli as in the study of Melnik et al. were used. An alternating contrast checkerboard was presented on a monitor (visual angle: $3.5x3.5$) with a 12 Hz frequency. A $4x4$ pixel cross in the center of the checkerboard served as the fixation cross. This cross was also shown separately for 500 ms prior to the experiment to make sure that the participants focus their attention on the middle of the screen. The subjects were instructed to passively observe. Four blocks with a total of 1200 alternations were shown.

2.3.2 MP

Again, the same setting as in the study of Melnik et al. was used (Melnik et al., 2017). The participants were instructed to close their eyes. A sequence of ten beeps was presented with a constant pace of one per second. After that, they were to continue pressing the key "Arrow Down" with the right index finger at the same pace for about one and a half minutes. A beep after 90 key presses served as a signal for the participants indicating that the paradigm is over. The participants had to perform this experiment three times such that a total of 270 key presses was recorded for each subject.

2.3.3 vDM: N240

The setting of this paradigm differed a bit more from the one of Melnik et al.. The stimuli stayed the same being text labels of German color names on a monitor (visual angle: $1.5x0.6$). The same four text labels *rot* (red), *grün* (green), *blau* (blue) and *gelb* (yellow) were used in a ratio of 1:1:1:1. The sequence, in which the text labels were shown, was random. The participants were told to react only to the target stimulus *rot* with pressing the key "Arrow Down" with the right index finger.
It was planned to show the stimuli for 280 ms with an interstimulus interval of one second. Unfortunately, due to some problems with the parallelization of the recording with the *Traumschreiber* and the generation of the stimuli, these timings differed throughout the experiments. Some stimuli lasted longer whereas others were only presented shortly. Moreover, in contrast to the paradigm of Melnik et al. no online feedback was given when an error occurred (false press, missed press) and the interstimulus interval did not increase after an error. These features could not be implemented due to the inaccurate timings.

120 stimuli were presented in each of the four blocks such that a total of 120 target and 360 distractor stimuli were shown per subject.

2.4 Procedure

The participants were told to wash their hair prior to the experiment. This is done to reduce fatty oils on the scalp which reduce the conductivity. A cap size was chosen and the EasyCap was placed on the head such that the Cz electrode was in the middle of the head. Alcohol was applied to the skin using Q-tips to remove remaining oils. Ten ring electrodes were connected to the cap and the *Traumschreiber* as described beforehand. A conductive gel was applied to establish a connection between the scalp and the electrodes.
A liveplot of the EEG data served as a mean to assess the connectivity of the electrodes (see figure 2.4). Bad electrodes were improved until a clear signal was visible.

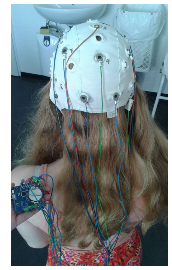


Figure 2.3: Subject with setup

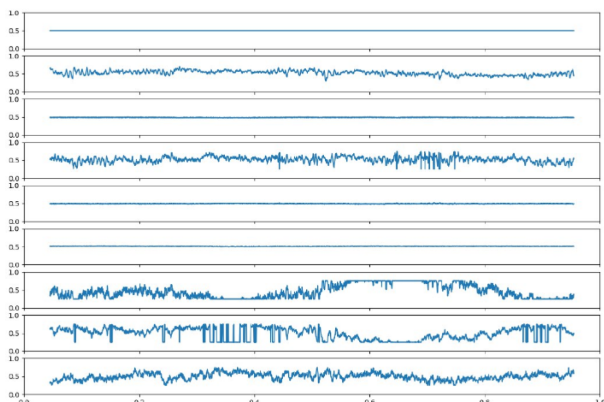


Figure 2.4: Liveplot during channel adjustment - The eight channels and the ground are displayed showing a liveplot of the activation. As can be seen, most channels are not connected properly. The second and the last channel show an activation pattern that is desired. This was used as a replacement of impedance measures.

2.5 Analysis

Preprocessing of data is an important step in data mining. It includes organizing the data, cleaning and modifying it (Cohen, 2014, p. 73). Since these steps might introduce some bias, all data was preprocessed in the same way. The danger of losing information and not having enough data to generalize are concerns to think of when choosing what amount of preprocessing to use (Blankertz et al., 2002). Considering this, the data was preprocessed in a way to allow analysis but avoid unnecessary preprocessing.

Since interfering artifacts often have higher frequencies than the signal of interest, it is common in ERP studies to apply filters. Filters like high or low pass filters, notch filters, bandstop filters all serve to increase the signal-to-noise ratio. No filter was applied in this study, as it seems to be unimportant for averaging because of the random appearance of noise. Single-trial analysis was also performed with the raw EEG data in order to keep the preprocessing as simple as possible.

2.5.1 Re-referencing

The choice of the reference is a highly discussed topic with no common ground. This choice depends on the number of electrodes used, their placement, the examined potentials and the applied analysis (Cohen, 2014, p. 82).
The *Traumschreiber* was built in a way to be maximally robust in case an electrode gets detached during sleep. Furthermore, its original purpose is to measure eye movement and ECG. For these measurements bipolar references are often used (Cohen, 2014, p. 82). Bipolar electrodes measure the difference between a pair of electrodes. The *Traumschreiber* uses nine bipolar electrodes to record eight channels and one electrode as the ground. But that is not the standard design for EEG systems.
Standard EEG systems use one reference electrode or the average of all electrodes is chosen as the reference. The problem of referencing to only one electrode is that if that one has a bad connection, all recorded data will be noisy. Moreover, the placement of that specific electrode has to be chosen thoughtfully because it affects the recorded data (Cohen, 2014, p. 17). Common choices for the reference electrode are averaged mastoids or earlobes (Cohen, 2014, p. 82). These electrodes have the advantage of being close to the brain but at the same time record nearly no brain activity. Cohen suggests that firstly, the reference electrode should not be placed on one hemisphere since the recorded data may otherwise have a lateralization bias. Secondly, the reference should not be near the origin of the assumed potential (Cohen, 2014, p. 82). If many electrodes are used, an average reference is recommended.
Since re-referencing is only a linear transformation of the data, it can be done afterwards and will not destroy any information. Therefore, the reference during recording is not as important, as it can be changed afterwards.

The procedure for re-referencing the data acquired with the *Traumschreiber* is not computationally costly. The resulting channels can be seen in figure 2.5 when a re-referencing to channel 4 is performed. As the *Traumschreiber* uses nine electrodes, data of nine channels should be recorded. Thus, a new channel is introduced. The cumulative sum over all channels is computed and the channel which should be the reference is subtracted. This procedure leads to nine channels where one is equal to zero since it displays the difference from the reference electrode to itself. That channel can be excluded for further analysis. Due to the calculation of the cumulative sum, the noise also adds up and increases the farther the electrode is from the reference. That is why a reference near the originating signal was chosen for each paradigm respectively. Even though Cohen claims that such an electrode is a bad choice as a reference (Cohen, 2014, p. 82), for the sake of this experiment it is the right choice. This is because this study is not interested in localizing the examined potentials. Instead, it is of interest to find good placements of electrodes to train a BCI with. Choosing an

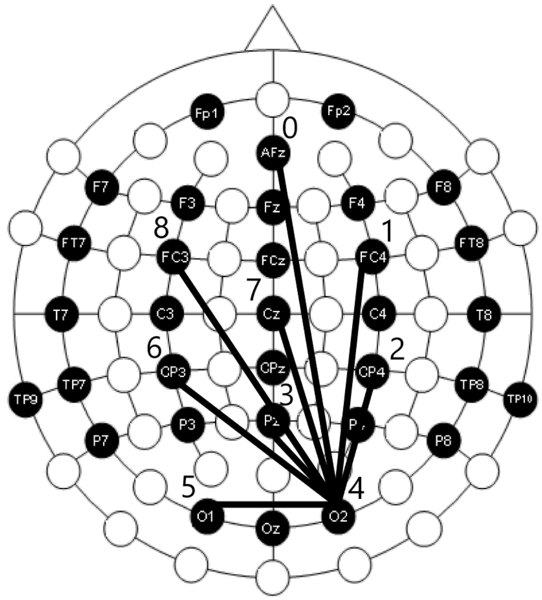
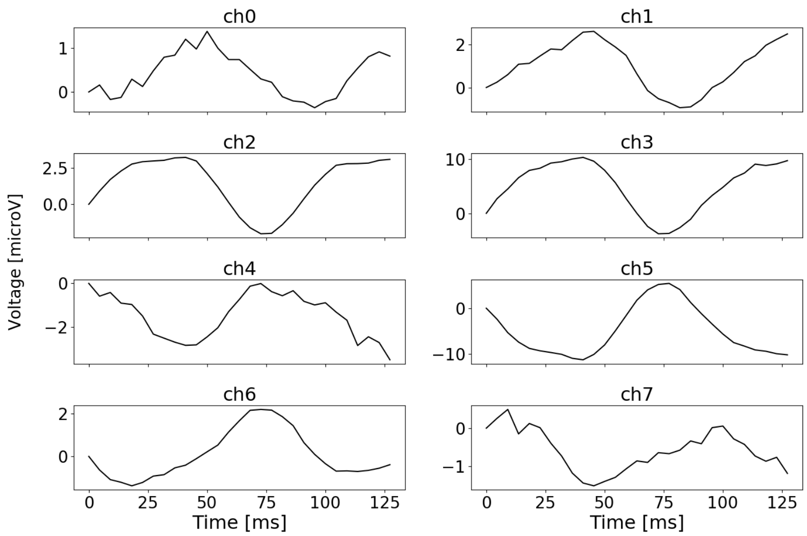
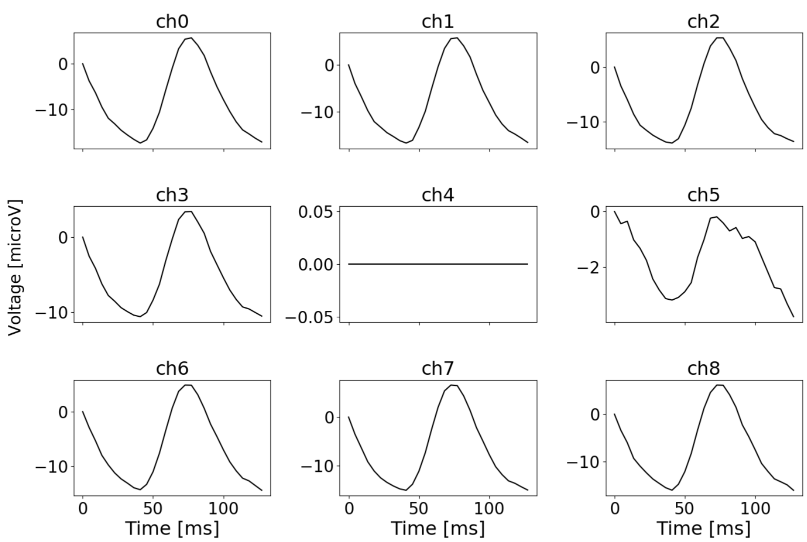


Figure 2.5: Channels after re-referencing to electrode four. Adapted from (FieldTrip, 2018).

electrode above the originating region as the reference allows an analysis of various different reference placements. That is because only one recording electrode is used, the one above the originating area, and the other electrodes serve as references. This allows to determine whether the reference placement influences the recorded data strongly or not.
For each paradigm a reference electrode was chosen out of the ones above the originating areas by plotting the ERP after re-referencing to each electrode. The electrode that showed a clear ERP in most of the channels was chosen as the reference. For SSVEP the electrode 4 (O2) was chosen, for N240 electrode 5 (O1) and for MP electrode 8 (FC3) was chosen.
The not re-referenced data itself already showed the ERP in some channels. The SSVEP showed in several channels and quite strongly in channel 3 and 5. Those channels each included one electrode that was placed above the visual cortex. The ERPs calculated with the not re-referenced data were on average smaller than the re-referenced ones (see figure 2.6). Re-referencing enhanced the signal in all channels since the reference electrode was chosen to be above the origin of the potential. As can be seen in figure 2.6 two channels remained the same after re-referencing. The potential of channel 3 is only inverted and the potential of channel 4 can be found in channel 5.



(a) Not Re-referenced



(b) Re-referenced to electrode 4

Figure 2.6: Before and After Re-referencing

2.5.2 Epoching

In order to extract those data points that are of interest, the continuous data is chunked into epochs. To do so, the markers that indicate the specific events are used. An epoch normally starts with the onset of the stimulus and lasts a few hundred milliseconds up to a second. For the SSVEP component one alternation is included in an epoch which lasts for 132 ms. The N240 component shows about 240 ms after the stimulus onset, so 300 ms after the onset were selected. For the motor potential, 50 ms before the tap and 250 ms after the tap form one epoch.

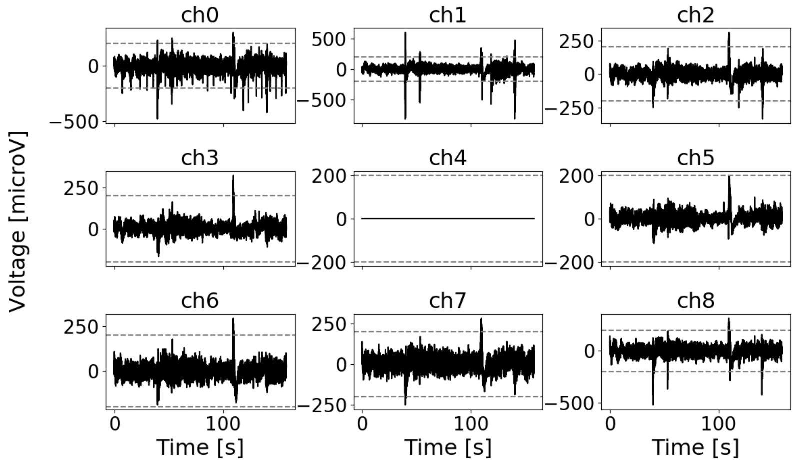
Since the *Traumschreiber* does not have a constant sampling rate and the transfer via Bluetooth also buffers the data, sometimes less data points arrived in a second and thus some extracted epochs had less data points. Those short epochs had to be excluded for further analysis. Thus, less epochs were extracted than stimuli were shown. For the SSVEP component 1.25 epochs were excluded on average per subject. For the N240 3.875 targets were excluded and 12.375 distractors. For the motor paradigm the data of about 22.75 taps was removed.

Reasons why that many trials had to be excluded are that the used timing of pygame is not that precise. Moreover, for the N240 paradigm there was once a connection failure and the *Traumschreiber* did not record that many data points for one block of one subject. For the N240 component erroneous epochs were also excluded. There has been a problem with the synchronisation of the data storage and the key presses. This may be the cause for the loss of that many trials of the motor task.

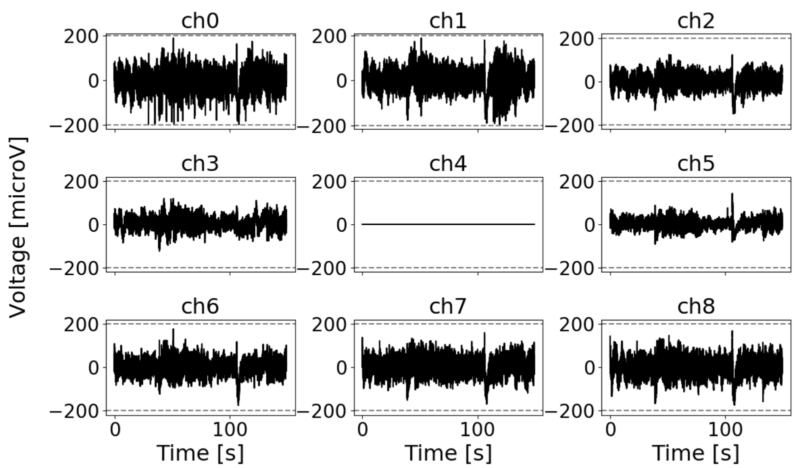
2.5.3 Artifact Removal

A lot of other signals that are way larger than the EEG signal are picked up during recording. Eye blinks and muscle movement can conceal the data as they produce stronger potentials. Other causes can also result in artifacts like touching an electrode or sweating. Since these potentials are not of interest, one way to deal with them is to remove them. Trial rejection by hand and automatic removal are both applied in EEG research. On the one hand automatic removal has the advantage of being fast and unbiased but also suffers from Type 1 and Type 2 errors (Cohen, 2014, p. 81). On the other hand manual removal is time consuming and relies on the visual inspection of the researcher.

For this study, artifacts mainly showed in form of drifts and high peaks as can be seen in figure 2.7. Those sharp peaks reflect artifacts resulting from either touching an electrode or body movement that affected the cables of the electrodes. In order to eliminate these an automatic removal was applied. Epochs with values higher than 200 microV in any channel were discarded. For each subject on average 61.625 epochs were excluded from the SSVEP data, 6 targets and 19.25 distractors from the N240 data and 12.125 taps from the motor data were removed because of artifacts.



(a) Before



(b) After

Figure 2.7: Before and After Artifact Removal

2.5.4 Baseline Correction

In order to reduce the variance in the data, baseline correction can be used. This shifts the potential on the y-axis such that at the start of the epoch the signal always starts at 0 microV (see figure 2.8). In this study, this was achieved by subtracting the first value of each epoch from the other values of that very epoch. Averaging and logistic regression benefit from this preprocessing step because the potentials are more aligned.

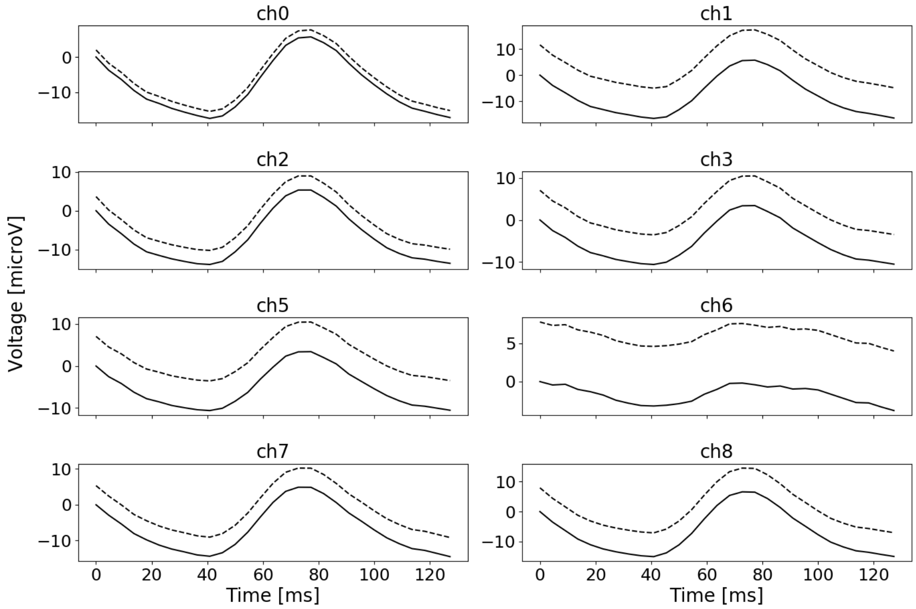


Figure 2.8: Before and After Baseline Correction

2.5.5 Logistic Regression

In order to implement the logistic regression model, the sklearn package in python was used. A logistic regression model with L2 regularization was trained. To find suitable values for the L2 regularization parameter the sklearn GridSearchCV was used. A grid search was performed on 7 parameters between $C = [10^{-3}, 10^{3}]$. A parameter was chosen for each model respectively. The variances of the two classes were kept similar by having about the same number of representatives for each class.
The input data of the class "0" was the preprocessed EEG data. Those examples were labelled as "positive data". The "negative data" forming the class "1" was randomly drawn from the data of all subjects and all paradigms and preprocessed in the same way as the positive data set. The data of the two classes were split into a training and a test set by using 0.25 of the data for testing. The shape of the input vector varied according to the respective paradigm and the applied method. The data of each paradigm was epoched using a different length. Therefore with a sampling rate of 220 Hz, the input vector has a shape of $(\frac{1}{220} * \text{length of epoch}) * (\text{number of channels}) x1$. The size of the input vector determines the number of parameters that need to be fitted. Since the extracted epochs for SSVEP are smaller than the ones for MP and N240, less parameters need to be fitted for this paradigm.

In order to evaluate the performance of the different models and to compare them, the AUC will be used. The receiver-operator characteristic (ROC) plot illustrates both the sensitivity, true positives and the specificity, true negatives of a model (Krzanowski & Hand, 2009). The AUC reflects a perfect prediction with a score of 1 and purely random chance prediction with a score of 0.5. Thus, AUC was chosen as a metric considering classifier specificity and sensitivity.

As BCIs focus on optimizing the application for one subject, a logistic regression model is fitted for each subject individually. Three different approaches to this will be performed.
Firstly, for each channel a model is fitted. The best channel can be determined by using the AUC as a metric. If even one channel is enough to train a model with a high accuracy, it can be concluded that BCIs only need one channel in order to classify an ERP. If the resulting AUC scores do not differ a lot between channels, it can be concluded that the placement of the reference is not crucial to get a good signal.
Secondly, the best three channels will be combined using them as additional parameters. This increases the amount of parameters that need to be fitted. Thus, the input vector's shape is now $(\frac{1}{220} * \text{length of epoch}) * 3x1$. This analysis serves to see whether more information of several channels improves the AUC score and would outperform a single channel BCI.
Thirdly, the best three channels will be combined without keeping the information that those signals are recorded from different channels. Thus, the input vector has a shape of $(\frac{1}{220} * \text{length of epoch})x1$. If this classification still produces scores way above the random chance prediction, one can conclude, that the electrode placement is not important because the classifier can learn with one placement and can still make good predictions using a different channel. Hence, the classifier of a BCI could be learned using one placement and afterwards electrodes could be placed somewhere else and the classification would still work. This could also be a solution for having only a short training phase and therefore only a small set of training data because several channels record similar information at the same time.

A further analysis will be done by training a model with the data of all subjects combined. This analysis shows whether combining the data of subjects could be a solution for having a small amount of training data. As single-trial analysis normally does not have to deal with variance across subjects, this will probably result in worse AUC scores. But perhaps the logistic regression model is able to find the underlying ERP signal like the traditional averaging method does. As each subject differs in strength of the potential and reaction time, this is rather difficult. This analysis can be used to compare the averaging method with logistic regression. If the model is able to perform well on data sets of various subjects, the learned coefficients should resemble the ERP that can be found by averaging. The coefficients should converge to the features that distinguish the two classes.

3 Results

The results show that the *Traumschreiber* is capable of recording prominent ERPs. For all paradigms the standard method of averaging showed the examined ERPs as described by Melnik et al.. Even without filtering, the results are satisfying. But due to bad impedances or an unstable setup, quite some data had to be rejected during epoching or because of artifacts. Moreover, the collection of data was interrupted early because the switch of the *Traumschreiber* fell off during the experiment with subject 9.
For a first analysis the traditional approach will be shown. The average of each channel and the average across all channels will be presented. Because of the way the re-referencing was performed, the ERP showed in most of the channels (see figures 3.2, 3.5, 3.9, 3.10 and 3.8). The results of the single-trial analysis using logistic regression will be shown for each paradigm. The real BCI use case will be examined first using the three different approaches as explained in the previous chapter. Thereafter, the further analysis using the data of all subjects will follow.

3.1 SSVEP

Figure 3.1 shows the variance across subjects for the SSVEP. This plot makes clear that the response of all subjects differed a bit, but especially the response of subject 9 is different from the others. In response to a flickering stimulus, the brain does not only flicker at the exact same frequency but also in integer harmonics of it. This explains the different appearance of the potential for subject 9. For the following analysis where subjects are combined subject 9 was left out, in order to ensure equal potentials. Subject 9 was included for the single-trial analyses that were performed for each subject individually in order to show that BCIs are capable of learning such a behaviour as well.

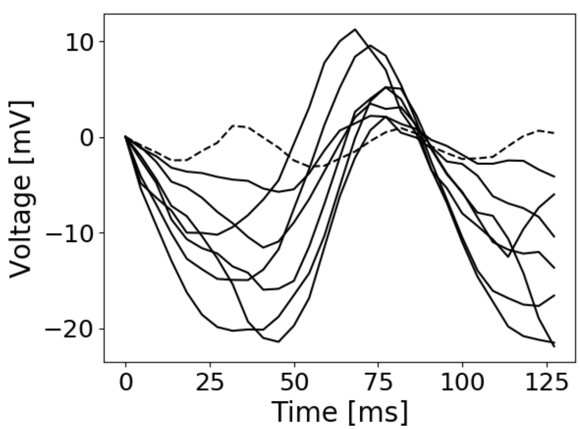
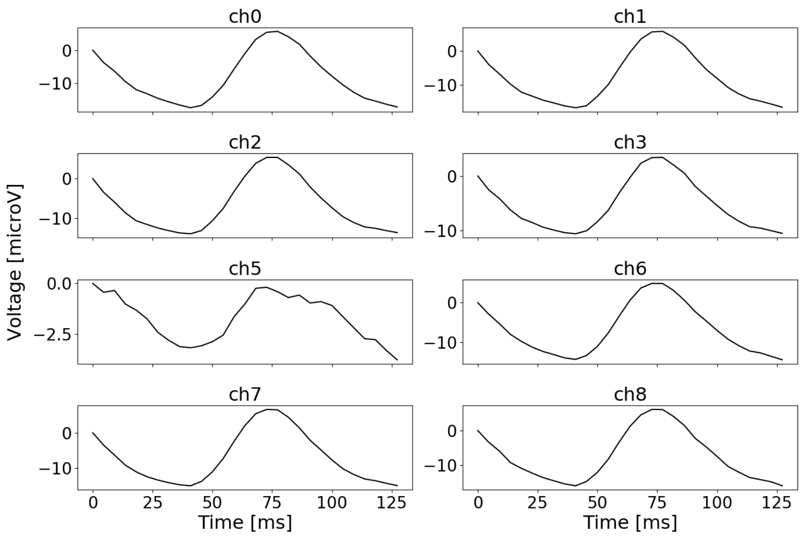
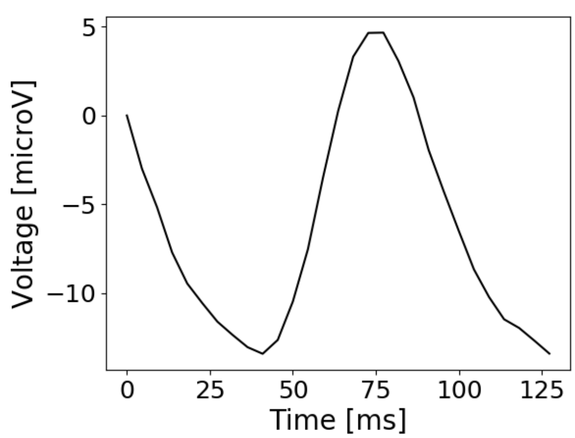


Figure 3.1: SSVEP: variance across subjects

Using the standard method of averaging, SSVEPs show in each channel and the average over all channels looks like the one observed by Melnik et al. (see figure 3.2).



(a) SSVEP for each channel



(b) SSVEP over channels

Figure 3.2: SSVEP: Averaging

Firstly, the results of the subject specific analysis will be presented. The AUC for each channel can be seen in table 3.1. A quarter of the data points of class "0" were used for testing. A mean AUC across subjects and channels of 0.685 was found. The mean AUC of the single best channel across subjects is 0.739. Table A.1 in the appendix shows all AUC scores of all methods and the calculated means. The location of the best channel differed across all subjects. But as the reference was chosen to be at the O2 electrode, all channels recorded SSVEPs and were able to provide good classifications (see table 3.1). The best channel across subjects is channel 1, FC4 with an AUC of 0.718. As can be seen in table 3.1, channel 5 has the lowest AUC scores across all subjects (0.602). Channel 3 and 5 display the same data as before re-referencing being the difference between O2 and Pz and O2 and O1, respectively. The not re-referenced data should have less noise in them, but in case of channel 5 the recording electrode and the electrode O1 were both placed above the originating region. This may have caused the underlying potential to vanish. Since Pz is not above the originating region, the AUC score for channel 3 is quite good (0.642), although not as good as the other channels (all above 0.68). This may be because Pz is still located close to the visual cortex and thereby can also record data of the visual cortex. The data of all channels other than those that were placed above or near the visual cortex as well was good enough to achieve AUC scores far above random prediction. The placement of the electrode seems to be not as important, as all channels produced comparable results unless both electrodes were located near or above the visual cortex.

channel subject	0	1	2	3	5	6	7	8
1	0.789892	0.760029	0.712143	0.722733	0.601468	0.760029	0.733229	0.743221
2	0.607308	0.650908	0.636707	0.583001	0.565509	0.650785	0.632993	0.596616
4	0.758368	0.756638	0.732226	0.682036	0.651401	0.731003	0.757018	0.743368
5	0.668265	0.687458	0.703304	0.713756	0.562684	0.715745	0.680428	0.650652
6	0.665390	0.715876	0.744021	0.576429	0.578668	0.679840	0.692760	0.681281
7	0.823851	0.821269	0.816025	0.661478	0.555633	0.804954	0.791178	0.806161
8	0.731557	0.727459	0.729508	0.631148	0.643443	0.754098	0.723361	0.713115
9	0.600929	0.627620	0.672044	0.568786	0.657647	0.580240	0.539241	0.576377
mean	0.705695	0.718407	0.718247	0.642421	0.602056	0.709587	0.693776	0.688849

Table 3.1: SSVEP: AUC for each subject and channel

Selecting the best three channels for each subject based on the AUC score and combining them, such that three times as many parameters were used, led to a mean AUC score of 0.720 (see table 3.2). This score is lower than the one obtained by using only the single best channel for each subject. Thus, combining the most informative channels in such a way does not improve the performance compared to using only the best channel in this case. This might be because each channel recorded the same signal as only the reference differed. But still, BCIs using only one channel can in some cases outperform ones that use several electrodes.

Using the data of the selected channels disregarding that the data was recorded from different channels increases the amount of data points that can be used for training. In this case, a lot of data points were available for training so adding more data was not so profitable. And as this adds more variance, a lower mean AUC across subjects of 0.724 than the mean of the single best channels was found (see table 3.2). As the score remained comparably high even though more variance was added, the different channels must have recorded similar signals. Thus, all those channels that led to the highest AUCs are good choices for electrode placements when recording SSVEPs.

subject	best 3 channels as params	best 3 channels as one
1	0.764063	0.766376
2	0.622468	0.639024
4	0.736067	0.730020
5	0.710241	0.736670
6	0.698255	0.701771
7	0.812362	0.820327
8	0.758197	0.761532
9	0.657344	0.639744
mean	0.719875	0.724433

Table 3.2: SSVEP: AUC for selected channels

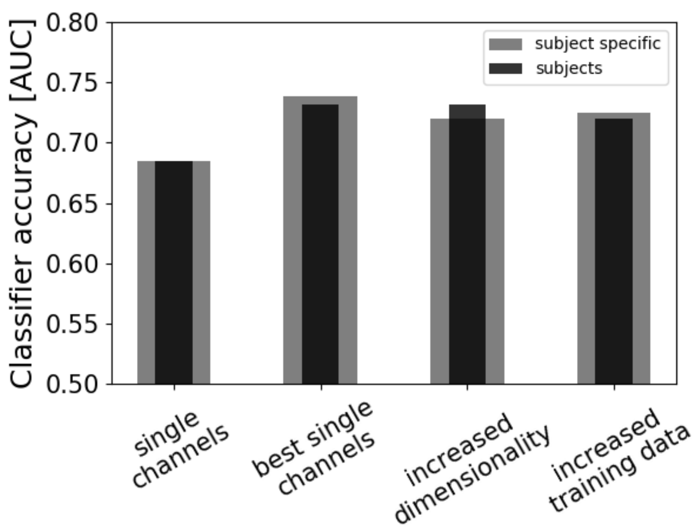


Figure 3.3: SSVEP: Comparison of methods

To sum up the subject specific analysis, the best AUC scores were achieved by using the single best channel (see figure 3.3). But also the different combinations of the best channels were able to produce comparably good results. This indicates that the choice of the reference electrode is not that important for BCI systems. Even the data of different channels can be used to train one model that can later perform classifications for a different channel.

A challenge for BCIs is that there are often not that many examples for training. A way to deal with that is to use additional data collected from other subjects. As the underlying ERP should be similar, data from other subjects can also provide informative data for the model. The same analyses as shown beforehand were performed for all subjects combined (excluding subject 9) to see whether this can be an option (see figure 3.3). The results can be seen in the appendix A.1. Remarkably good results were obtained for all channels and also the combinations of the best three channels yielded good AUC scores around 0.7. The best score was achieved by channel 2, CP4 with an AUC of 0.732. When using that much data, increasing the amount of parameters led to better results than beforehand. All in all, the results are comparable to the ones of the subject specific analysis. Those results indicate that the SSVEP is quite similar across subjects and thus a model trained with data of other subjects can still make good predictions for a different subject. It also shows that logistic regression is capable of dealing with variance when regularization is applied.
Combining all subjects reflects what the standard averaging method does. As a logistic regression model learns coefficients that allow to discriminate the two classes, it should converge to something that resembles the underlying ERP. A logistic regression using the data of all subjects (except subject 9) and channels was trained. The learned coefficients were scaled to match the demeaned ERP across channels. The coefficients curve is smooth and has a similar shape like the averaged ERP across channels and subjects (see figure 3.4). This shows that a logistic regression model is able to learn the underlying ERP.

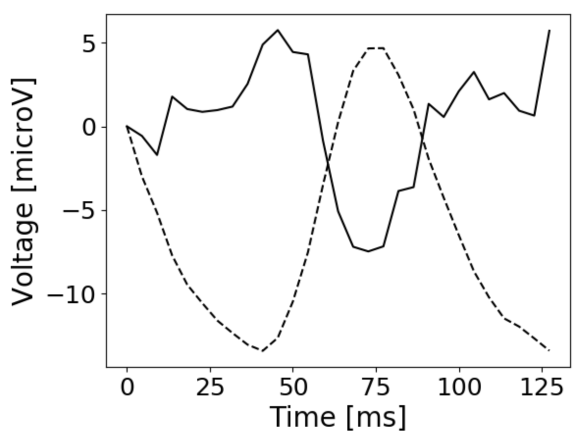
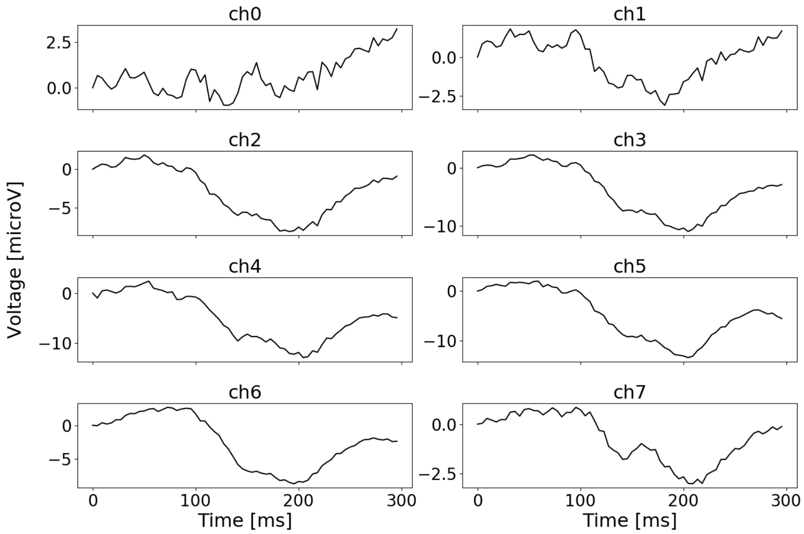


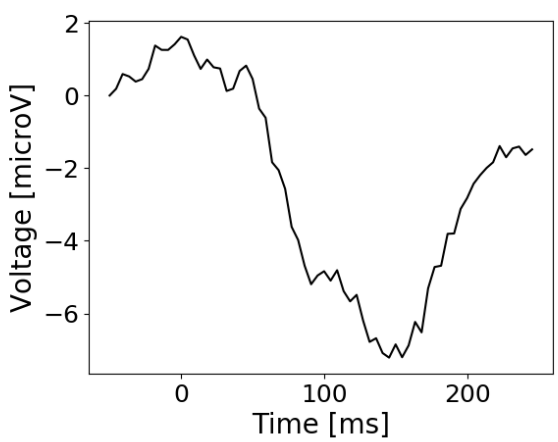
Figure 3.4: SSVEP: Comparison Averaging Logistic Regression Coefficients

3.2 MP

The standard method of averaging produced good results in most channels and also the average over all channels looks like the one described in the literature and the one observed by Melnik et al. (see figure 3.5). The channels 0 and 1 did not record the MP that well. This may be due to the fact that the electrodes AFz and FC4 are also located near or above the motor cortex and hence also recorded the ERP like the reference electrode at FC3.



(a) MP for each channel



(b) MP over channels

Figure 3.5: MP: Averaging

The subject specific analysis achieved a mean AUC across all subjects and channels of 0.609. The mean across the single best channel was 0.693. This even includes those channels that already showed a bad average in the traditional analysis. Channel 0 and 1 had the lower AUC scores like it was already assumed based on the averaged ERPs (see figure 3.5). But also channel 7, the one that displays the not re-referenced data, had low scores which can be explained by the location of the electrode Cz being near the originating area. Those three channels obtained a mean score below 0.6. The other 5 channels all achieved a mean AUC

channel subject	0	1	2	3	4	5	6	7
1	0.537193	0.634208	0.660809	0.670077	0.686206	0.652744	0.685243	0.657198
2	0.557895	0.616667	0.614912	0.649123	0.632456	0.604825	0.739474	0.665351
4	0.451136	0.623214	0.640422	0.622240	0.632143	0.613312	0.746916	0.539773
5	0.586201	0.576299	0.550487	0.638636	0.694481	0.675812	0.639123	0.512825
6	0.542075	0.526457	0.588695	0.572844	0.573776	0.610956	0.647902	0.526807
7	0.620313	0.620313	0.677557	0.651562	0.594176	0.544886	0.663068	0.578551
8	0.535384	0.501984	0.592262	0.546296	0.583333	0.673280	0.516204	0.536045
9	0.506641	0.628874	0.641524	0.674889	0.668564	0.665718	0.575111	0.553289
mean	0.542105	0.591002	0.620833	0.628209	0.633142	0.630192	0.651630	0.571230

Table 3.3: MP: AUC for each subject and each channel

subject	best 3 channels as params	best 3 channels as one
1	0.653587	0.687290
2	0.655263	0.682357
4	0.539448	0.664487
5	0.576461	0.671514
6	0.495804	0.625337
7	0.605824	0.644341
8	0.552249	0.654211
9	0.631088	0.638939
mean	0.588715	0.658559

Table 3.4: MP: AUC scores for combined best channels of each subject

above 0.62. The best channel differed across the subjects, but channel 6, CP3 had the best mean AUC with 0.652.

For the MP only a few data points were recorded. After the preprocessing 235.125 epochs could be used as input vectors, so only 176.34 positive training examples were used on average. As the MP has 66 parameters, the accuracy may be improved by adding more data points. Adding more parameters by selecting the best three channels did not improve the AUC score. This was expected because using less positive training examples than parameters is not useful (Peduzzi et al., 1996). A mean score of 0.589 was obtained as can be seen in table 3.4.

Using data of the best three channels as if they were recorded by one channel seems to be more promising because it increases the amount of data and not the number of parameters. This method led to a mean score of 0.659. This is below the score the single best channels but way above random chance prediction.

The best results were achieved by using the single best channel but also the combination of the three best channels to increase the amount of data achieved good scores (see 3.6). Therefore, it can be advantageous to increase the training examples by using the input of several channels.

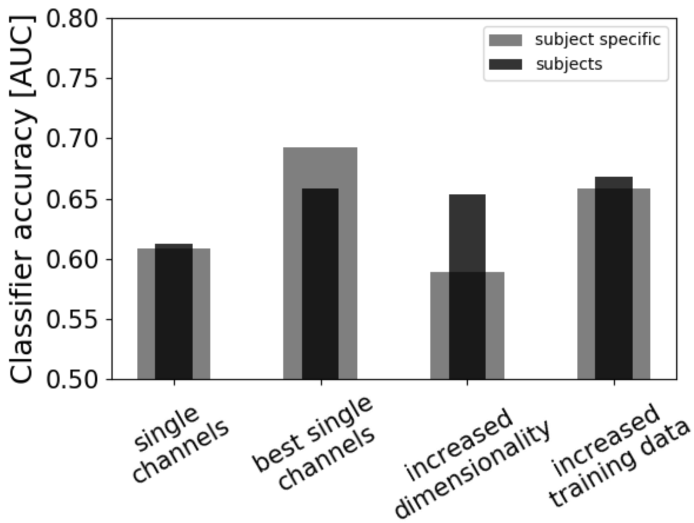


Figure 3.6: MP: Comparison of methods

Another way to add more positive examples is by using the data of several subjects. The results of this can be seen in the appendix in table A.2. The results were comparable to the ones of the subject specific analysis and for some methods even better (see figure 3.6). The best score with an AUC of 0.658 was achieved by channel 3, Pz. This is even higher than the score of the common best channel across subjects (0.652). As more data points were used, both combinations of the best three channels, resulted in better scores than beforehand (increased dimensionality: 0.589 vs 0.653, increased training data: 0.659 vs 0.668). This shows that using data of other subjects to increase the data for paradigms where only a few data points are available for training helps to increase the accuracy of the classifier. Even though this adds variance, it helps to increase the AUC, especially when there are a lot of parameters to fit.
The comparison of the average with the learned coefficients where both methods used the same data can be seen in figure 3.7. The curves to not resemble each other strongly. The coefficients curve is not smooth. The model did not assign high values to those parameters that should be the most informative features. Those are the positivity at 0 ms and the negativity at 150 ms.

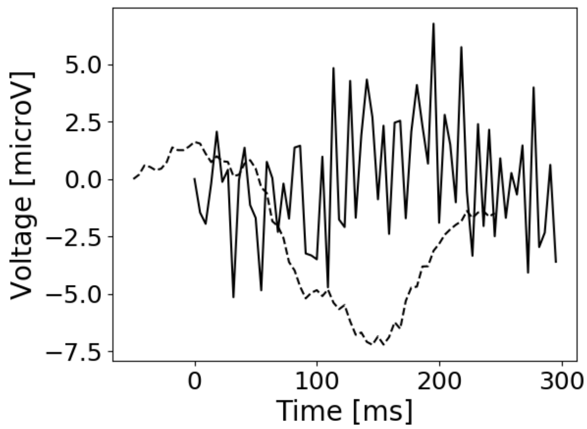
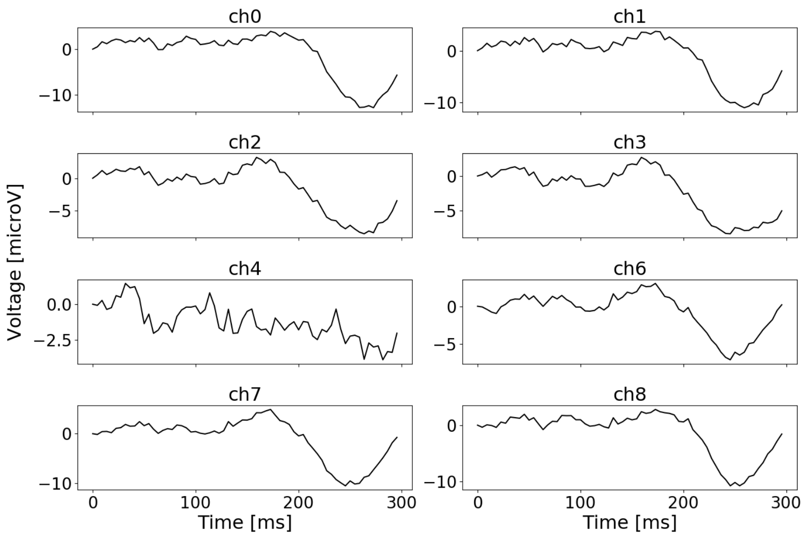


Figure 3.7: MP: Comparison Averaging Logistic Regression Coefficients

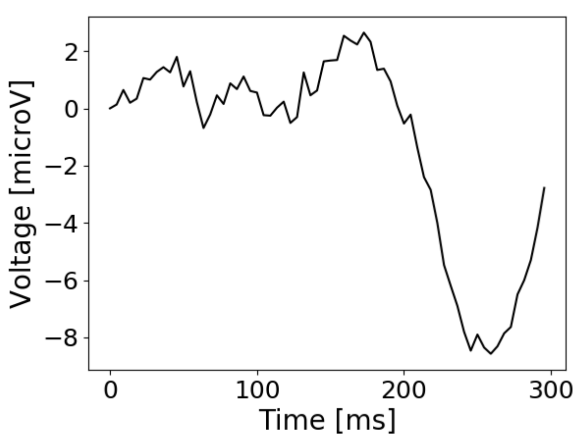
3.3 N240

In case of the N240 component the grand average is often calculated. This is the difference between the target and the distractor ERP (Melnik et al., 2017). Again the ERP showed in most of the channels and only the ERP of channel 4 is less clear. An explanation for this can be that O1 and O2, the reference are both located above the visual cortex. They may have picked up a similar signal and thus their difference does not show that activation. Channel 4 and 6 display the same potential as before re-referencing. The average over the channels is comparable to the results that Melnik et al. obtained in their study. The negativity is located around 240 ms after the stimulus onset which was expected.

As the logistic regression model will be trained using the data of the targets and distractors separately, the average of the target ERP and the distractor ERP was calculated as well and can be seen in the figures 3.9 and 3.10. For the target ERP the signal could be observed in all channels although they do not look exactly the same. The averaging method led to good results for the N240 distractors as the potential showed in most of the channels clearly (see figure 3.10). The curves differ across channels but still resemble each other.

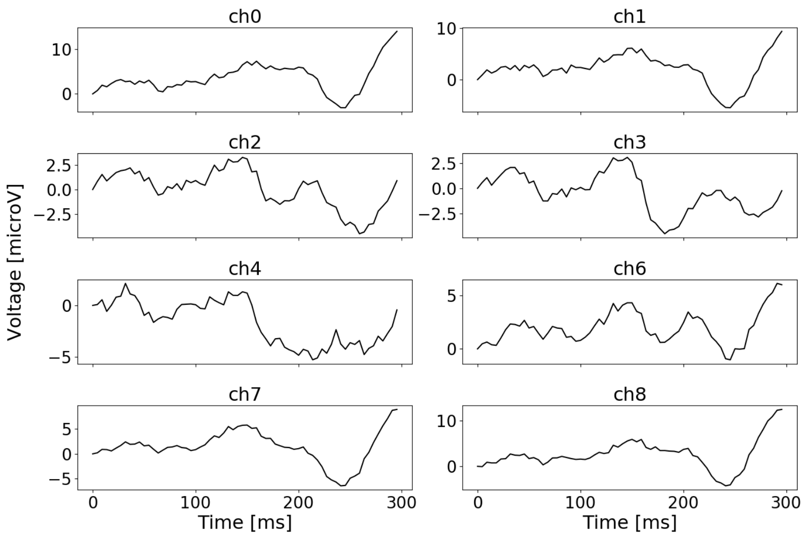


(a) N240 target minus distractor for each channel

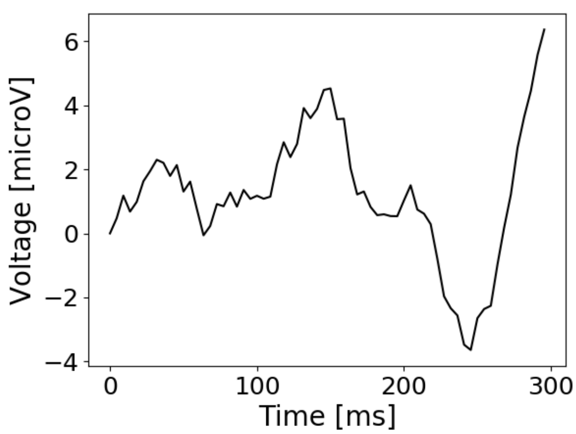


(b) N240 target minus distractor over channels

Figure 3.8: N240: grand average

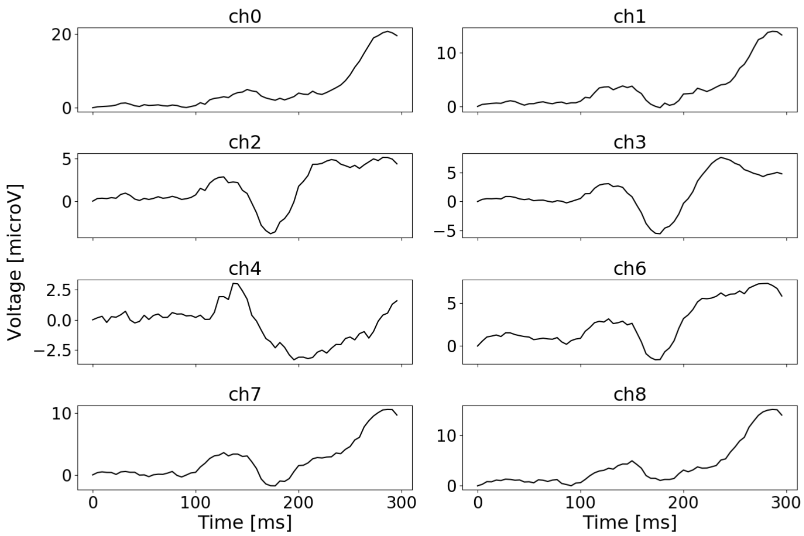


(a) N240 target for each channel

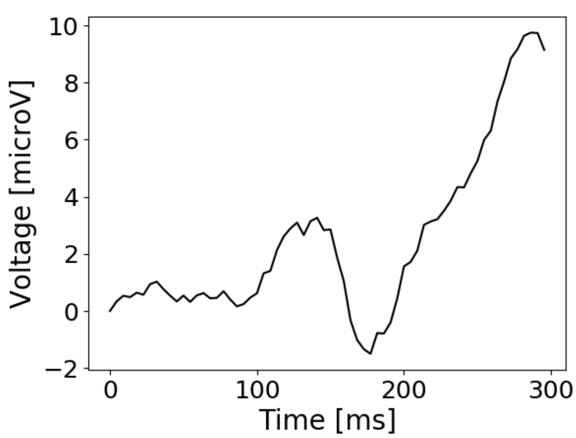


(b) N240 target over channels

Figure 3.9: N240 target: Averaging



(a) N240 distractor for each channel



(b) N240 distractor over channels

Figure 3.10: N240 distractor: Averaging

The channel that got the highest AUC scores across subjects for the N240 target component is channel 1, AFz (see table 3.5). But all channels performed well with a mean score across channels and subjects of 0.655. Only channel 4, O2 had a score below 0.6 which is again due to its location above the visual cortex. The single best channels got a mean AUC of 0.757.

channel subject	0	1	2	3	4	6	7	8
1	0.852163	0.871394	0.700721	0.673077	0.712740	0.653846	0.735577	0.704327
2	0.771969	0.719132	0.685762	0.720245	0.515573	0.606229	0.706340	0.749166
4	0.634038	0.700779	0.754727	0.568409	0.536151	0.622358	0.669633	0.600667
5	0.631010	0.681490	0.554087	0.557692	0.546875	0.667067	0.697115	0.700721
6	0.687500	0.847222	0.854167	0.625000	0.465278	0.694444	0.680556	0.784722
7	0.571429	0.553571	0.642857	0.607143	0.535714	0.571429	0.464286	0.589286
8	0.639752	0.593944	0.758540	0.574534	0.643634	0.689441	0.653727	0.554348
9	0.634118	0.659412	0.623529	0.541765	0.639412	0.703529	0.694118	0.628824
mean	0.677747	0.703368	0.696799	0.608483	0.574422	0.651043	0.662669	0.664008

Table 3.5: N240 target: AUC score for each subject and each channel

Each logistic regression model with 66 parameters was trained with about 84 positive training examples. Therefore adding more parameters will not improve the results. Fitting a model with 198 parameters using only 84 positive training examples led thus only to a mean AUC of 0.638 (see table 3.6). That score is still surprisingly high which may be due to the applied regularization.

As for the MP also the N240 target component would benefit from increasing the amount of training data. Combining the three best channels led to a mean AUC score of 0.756 (see table 3.6). This is higher than the mean of all single channels and comparable to the AUC of the single best channels 3.11. This again shows that the ERPs that were recorded are quite similar and in order to increase the amount of training data different channels can be used to record simultaneously a similar signal.

subject	best 3 channels as params	best 3 channels as one
1	0.680288	0.826738
2	0.641824	0.782883
4	0.586763	0.801200
5	0.618990	0.734826
6	0.694444	0.864311
7	0.535714	0.683233
8	0.701087	0.718814
9	0.644706	0.636264
mean	0.637977	0.756034

Table 3.6: N240 target: AUC scores for combined best channels of each subject

Again the mean of the single best channel and the score obtained by increasing the training data were the highest (see figure 3.11). As was already anticipated the increase in parameters did not improve the performance.

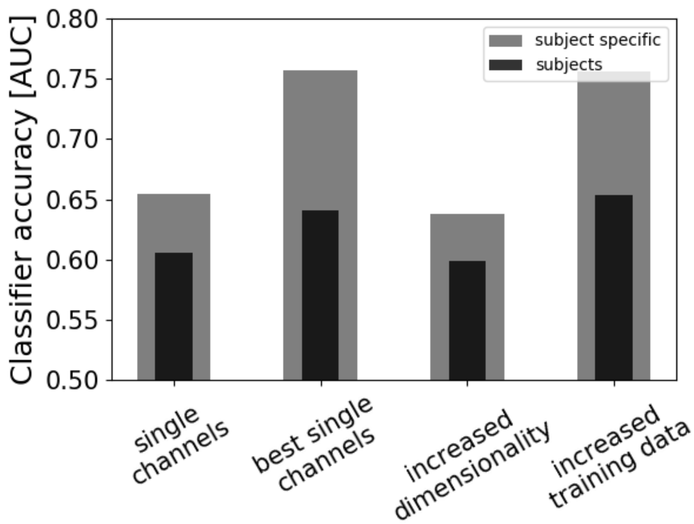


Figure 3.11: N240 target: Comparison of methods

Using the data of all subjects also increases the amount of data. But for the N240 target component the variance between subjects must have been quite high as the results were not that good. The combination of the best three channels also did not achieve a high AUC. Results around 0.6 were obtained which is far below the scores of the subject specific analysis (see figure 3.11). The exact AUC scores can be seen in the appendix in table A.3.
Due to the strong variance, the coefficients curve is not smooth and does not resemble the average that much (see figure 3.12). The bad performance of the logistic regression models using the data of all subjects already indicated this because it already showed that a lot of

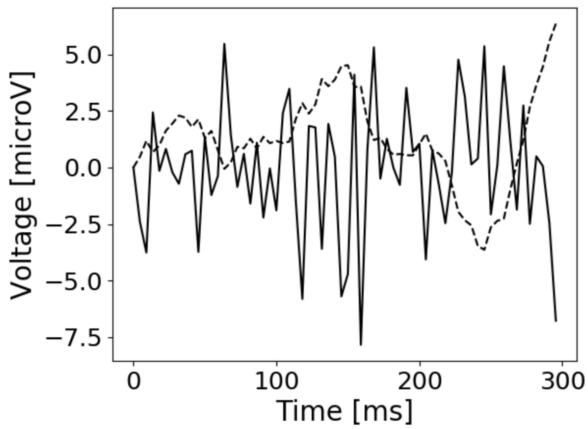


Figure 3.12: N240 target: Comparison Averaging Logistic Regression Coefficients

variance is in the data.

For the distractor component the best channel across subjects is also channel 1, FC4 (see table 3.7) with a mean AUC of 0.703. The mean of the single best channels is 0.757. The mean across all channels with a score 0.655 is still quite high but clearly below the mean of the single best channels. Here, channel 4 also had the lowest scores which can be explained with the reasons already mentioned beforehand.

channel subject	0	1	2	3	4	6	7	8
1	0.765547	0.765547	0.745231	0.686112	0.681866	0.716553	0.759733	0.748106
2	0.854545	0.780682	0.730556	0.758586	0.645707	0.752652	0.825758	0.797222
4	0.802285	0.881384	0.825013	0.779686	0.655771	0.768514	0.785432	0.785368
5	0.678152	0.618058	0.635322	0.769762	0.672201	0.636130	0.575375	0.623567
6	0.754399	0.689150	0.807674	0.781036	0.684995	0.745112	0.787146	0.791056
7	0.551034	0.649885	0.613103	0.681839	0.623678	0.670345	0.635172	0.603678
8	0.806420	0.710370	0.776543	0.748889	0.712346	0.673827	0.693333	0.721728
9	0.696970	0.543561	0.674373	0.680316	0.640413	0.657458	0.725575	0.611351
mean	0.677747	0.703368	0.696799	0.608483	0.574422	0.651043	0.662669	0.664008

Table 3.7: N240 distractor: AUC scores for each subject and each channel

Compared to the MP and N240 target component, the N240 distractor component has the most data points for fitting the 66 parameters. 246.28 positive training examples were available for fitting 198 parameters which is still a low ratio between events and parameters (Peduzzi et al., 1996). A mean AUC of 0.638 was achieved. This is below the score across all single channels and thus did not lead to an improvement.

Adding more data by using the best three channels got better results with a mean AUC of 0.761. This is even higher than the score of the single best channels (0.757). Again, the recorded signals must be really similar across channels and combining those can even increase the AUC. The placement of the electrodes is in this case again quite unimportant as it only needs to be taken care of that only one electrode is placed above the originating area.

subject	best 3 channels as params	best 3 channels as one
1	0.660766	0.809343
2	0.731944	0.839450
4	0.722357	0.821209
5	0.655451	0.634782
6	0.662757	0.872555
7	0.642529	0.628670
8	0.674074	0.805475
9	0.629310	0.678424
mean	0.637977	0.761239

Table 3.8: N240 distractor: AUC score for combined best channels of each subject

As can be seen in figure 3.13, the mean of the single best channels and the method of increasing the training data have the highest scores. Increasing the amount of parameters produced a lower score.

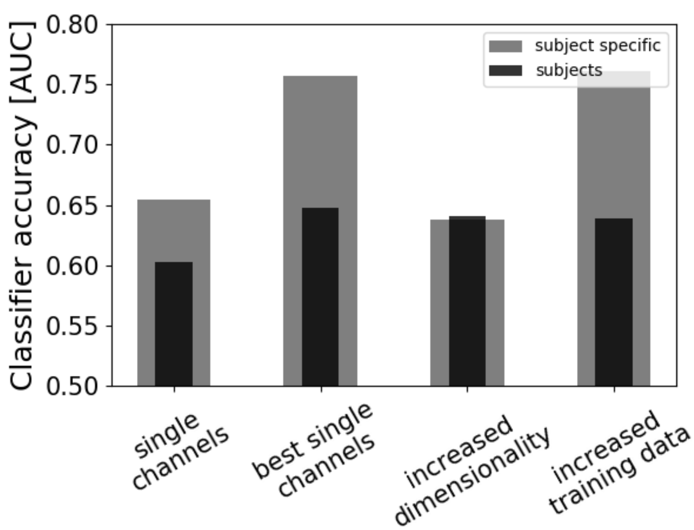


Figure 3.13: N240 distractor: Comparison of methods

Training a model with the data of all subjects did not result in high AUC scores (see table A.4). The AUC scores for the individual channels varied between 0.57 and 0.65. The same

was observed for the N240 target component. The results of the subject specific analysis are substantially better than the analysis using data of all subjects (see figure 3.13).
It appears that the N240 component differs across subjects more than for example the SSVEP component. This was also found by Melnik et al.. They reported that the variance across subjects was the largest source of variance and that this variance increased for more cognitive oriented tasks. Therefore, the combination of subjects works quite well for simple tasks like SSVEP and does not result in high AUC scores for the N240 component, a more complex task.
The average over all channels created by the standard method compared to the coefficients learned by a logistic regression that was trained with the data of all subjects and all channels can be found in figure 3.14. Due to the high variance the curve of the coefficients is not smooth and the potential as detected by the averaging method cannot be seen in the coefficients.

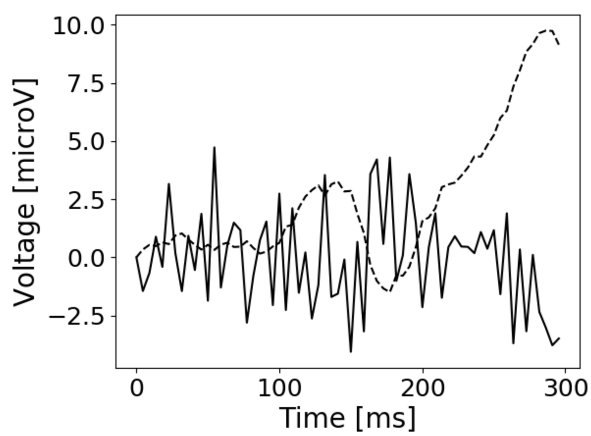


Figure 3.14: N240 distractor: Comparison Averaging Logistic Regression Coefficients

4 Discussion

This study shows that the data recorded with the *Traumschreiber* is sufficient to observe ERP components of different complexity. Concerns that are normally raised for mobile EEG systems concern the data quality, the sampling rate and non standard electrode locations (Krigolson et al., 2017). This study shows that the *Traumschreiber* which used only eight channels and a sampling rate of 220 Hz is able to perform traditional ERP analysis.
The data recorded with the *Traumschreiber* may not be useful for analyses that focus on the spatial location of ERPs. Since only eight electrodes are used and they were re-referenced in a way that only one location was examined, no information about the location of the potential can be obtained by this study. Most researchers use at least 32 electrodes to achieve a suitable spatial resolution (Cohen, 2014, p. 26).

4.1 Electrode Placement

The placement of the reference would ideally be common across studies as it can influence the data (Nunez et al., 1997). It was even found that two distinct ERP are actually the same and only were reported as distinct ones due to the applied reference placement (Joyce & Rossion, 2005). But as BCIs use well known ERPs with a known source of activity, the placement is not of much importance especially if the location of the potential is not used for analysis.
In this study nearly all channels and therefore different reference placements produced a high prediction score. The data was re-referenced in such a way that each channel measured the same signal that was recorded by the reference itself, the other electrodes served as the real references. It was found that the placement of the reference does not influence the performance of the classifier significantly unless it was placed near or above the originating area as well. In that case, both channels recorded the signal of interest and their difference did not show that potential. It can be concluded that for BCIs the placement of the reference does not matter significantly and that even a single channel can be used for classification.

4.2 Channel Selection

For the other approaches the three best channels were selected using the AUC as a metric. This step could be performed in a more advanced manner like using other machine learning techniques in order to find the best combination of channels (Arvaneh et al., 2010; Shan et al., 2015). As in this study most channels recorded the same signal, the different channels do not hold that much distinct information. Therefore combining those channels with the best AUC score also combines thóse with the highest predictive qualities.

4.3 Dimensionality

The combination of the best channels in a way to increase the amount of parameters did not produce scores far above random prediction preformance. This can be explained by the high number of parameters that needed to be fitted. Only for the SSVEP component, where more data points were available for training and fewer parameters had to be fitted, it produced

comparable results to the other methods. Therefore, this method seems to be only helpful if a lot of data is available and the extracted epoch is short, holding only informative features. In order to achieve good results for the other paradigms, shorter epochs could be extracted. This would reduce the amount of parameters and maybe exclude uninformative features. For example the N240 component shows its typical negative shift at around 240 ms after the stimulus onset. Extracting an epoch from 200 ms till 300 ms after the onset would decrease the amount of parameters that have to be fitted. This epoch would still contain the most informative features of the component.
Another way to reduce the dimensionality is to apply a different regularization. In this study, L2 regularization was chosen in order to avoid overfitting. Other researchers use a L1 regularization and thereby introduce sparsity (Shi, Xu, & Baraniuk, 2014; Tibshirani, 1996). This can be interpreted as adding an exponential prior, typically a Laplacian one (Shi et al., 2014). By doing so, it is assumed that only a subset of the features is informative. L1 regularization outperforms L2 regularization if there are many uninformative parameters (Ng, 2004). Using L1 regularization can be interesting for those paradigms where large epochs were extracted like MP and N240. Those epochs may include uninformative features and by adding sparsity the predictive accuracy could be improved and the classifier could be more robust.
Other studies combined the information of channels differently while applying the same idea. Instead of creating a large feature vector a matrix is used as the input containing information about time and channel. Bilinear logistic regression extends the classical logistic regression to enable it for two dimensional input (Dyrholm, Christoforou, & Parra, 2007). The resulting dimensionality reduction yields an improved generalization ability (Shi et al., 2014). Weight vectors along each dimension are learned. This extension was able to achieve high accuracies for BCI applications (Dyrholm et al., 2007; Stewart, Nuthmann, & Sanguinetti, 2014). In this study the channels recorded mainly the same signal due to the re-referencing. Therefore, combining those channels does not add that much information. But in general, this way of combining channels is a good idea for those potentials that can be measured at different locations.

4.4 Increasing Training Data

Facing another challenge of BCIs, the three best channels were combined as if they were recorded from one single channel. This can be a way to increase the amount of training data without increasing the training time of the subject. High scores were achieved that are comparable to the ones obtained by the single best channel. This shows that the channels recorded the same signal and hence the variance across channels was not that high.
Combining subjects is another way to increase the amount of training data without increasing the training time per subject. But this only worked out for low-level tasks which for example evoke the SSVEP component. More cognitive tasks like the N240 component vary significantly across subjects and therefore render the fitting harder. This was reported by Melnik et al. as they found that the variance across subjects is higher for those complex tasks (Melnik et al., 2017). But BCIs that use simple tasks could add other subjects' training data to a single subject's data and would achieve a high accuracy.

5 Conclusion

This study generally proved, that the low-cost EEG system, the *Traumschreiber*, is capable of recording ERPs and provides informative data for single-trial analysis. It also showed that questioning some standards of EEG research can help to improve BCIs.

It was found that the placement of electrodes is rather unimportant for BCIs. Well investigated potentials are used as the new communication and thus the recording electrode can be directly placed above the originating area. The reference should not be placed near that recording electrode as otherwise it might record the same signal.
It can be concluded that an increase of the dimensionality by combining the best channels is only useful if those channels hold distinct information. This was not really the case in this study due to the way the data was re-referenced.
Different approaches proved to be helpful to deal with the challenge of a small training data set. For all paradigms the data of several channels can be used to increase the training data. Using the data of different subjects proved only successful for low-level tasks as otherwise the variance across subjects is too high.

This study was able to show that the *Traumschreiber* can perform single-trial analysis and can thus be used as a BCI. Furthermore, it was shown that questioning standards in research can help to find solutions for modern application.

References

Anderson, J. R. (2015). *Cognitive psychology and its implications 8th edition.* Worth Publishers: A Macmillan Eductation Company.

Appel, K. (2017). *The traumschreiber system: Enabling crowd-based, automated, complex, polysomnographic sleep and dream experiments* (Unpublished doctoral dissertation). University of Osnabrück.

Arvaneh, M., Guan, C., Ang, K. K., & Quek, H. C. (2010). Eeg channel selection using decision tree in brain-computer interface. In *Proceedings of the second apsipa annual summit and conference* (pp. 225–230).

Bashashati, H., Ward, R. K., Birch, G. E., & Bashashati, A. (2015). Comparing different classifiers in sensory motor brain computer interfaces. *PloS one*, *10*(6), e0129435.

Berlin Brain - Computer Interface. (2018). *Bci competition.* Retrieved from `http://www.bbci.de/activities#competition` ([accessed 30-08-2018])

Blankertz, B., Curio, G., & Müller, K.-R. (2002). Classifying single trial eeg: Towards brain computer interfacing. In *Advances in neural information processing systems* (pp. 157–164).

Brain Products. (2018). *Easycap.* Retrieved from `https://www.brainproducts.com/productdetails.php?id=20` ([accessed: 27-08-2018])

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (2007). *Handbook of psychophysiology.* Cambridge University Press.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice.* MIT press.

Di Russo, F., Pitzalis, S., Aprile, T., Spitoni, G., Patria, F., Stella, A., ... Hillyard, S. A. (2007). Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential. *Human brain mapping*, *28*(4), 323–334.

Dyrholm, M., Christoforou, C., & Parra, L. C. (2007). Bilinear discriminant component analysis. *Journal of Machine Learning Research*, *8*(May), 1097–1111.

FieldTrip. (2018). *easycapm3 - extended 10/20-system with 30 channels.* Retrieved from `http://www.fieldtriptoolbox.org/_media/template/easycapm3.png` ([accessed: 17-09-2018])

Gao, X., Xu, D., Cheng, M., & Gao, S. (2003). A bci-based environmental controller for the motion-disabled. *IEEE Transactions on neural systems and rehabilitation engineering*, *11*(2), 137–140.

Herrmann, C. S. (2001). Human eeg responses to 1–100 hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental brain research*, *137*(3-4), 346–353.

Huang, Y., Erdogmus, D., Mathan, S., & Pavel, M. (2006). Boosting linear logistic regression for single trial erp detection in rapid serial visual presentation tasks. In *Engineering in medicine and biology society, 2006. embs'06. 28th annual international conference of the ieee* (pp. 3369–3372).

Joyce, C., & Rossion, B. (2005). The face-sensitive n170 and vpp components manifest the same brain processes: the effect of reference electrode site. *Clinical Neurophysiology*, *116*(11), 2613–2631.

Kaper, M., Meinicke, P., Grossekathoefer, U., Lingner, T., & Ritter, H. (2004). Bci competition 2003-data set iib: support vector machines for the p300 speller paradigm. *IEEE*

Transactions on Biomedical Engineering, *51*(6), 1073–1076.

Kappenman, E. S., & Luck, S. J. (2010). The effects of electrode impedance on data quality and statistical significance in erp recordings. *Psychophysiology*, *47*(5), 888–904.

Krigolson, O. E., Williams, C. C., Norton, A., Hassall, C. D., & Colino, F. L. (2017). Choosing muse: Validation of a low-cost, portable eeg system for erp research. *Frontiers in neuroscience*, *11*, 109.

Krzanowski, W. J., & Hand, D. J. (2009). *Roc curves for continuous data.* Chapman and Hall/CRC.

Lemm, S., Blankertz, B., Dickhaus, T., & Müller, K.-R. (2011). Introduction to machine learning for brain imaging. *Neuroimage*, *56*(2), 387–399.

Lemm, S., Schafer, C., & Curio, G. (2004). Bci competition 2003-data set iii: probabilistic modeling of sensorimotor/spl mu/rhythms for classification of imaginary hand movements. *IEEE Transactions on Biomedical Engineering*, *51*(6), 1077–1080.

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, *4*(2), R1.

Mandel, C., Lüth, T., Laue, T., Röfer, T., Gräser, A., & Krieg-Brückner, B. (2009). Navigating a smart wheelchair with a brain-computer interface interpreting steady-state visual evoked potentials. In *Intelligent robots and systems, 2009. iros 2009. ieee/rsj international conference on* (pp. 1118–1125).

Mandt, L. S. (2017). *Conducting a polysomnographic, automated auditory stimulation sleep experiment using a portable high-tech sleep mask* [Bachelor's Thesis].

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models.* Chapman & Hall.

Melnik, A., Legkov, P., Izdebski, K., Kärcher, S. M., Hairston, W. D., Ferris, D. P., & König, P. (2017). Systems, subjects, sessions: to what extent do these factors influence eeg data? *Frontiers in human neuroscience*, *11*, 150.

Ng, A. Y. (2004). Feature selection, l 1 vs. l 2 regularization, and rotational invariance. In *Proceedings of the twenty-first international conference on machine learning* (p. 78).

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: a review. *Journal of vision*, *15*(6), 4–4.

Nunez, P. L., Srinivasan, R., Westdorp, A. F., Wijesinghe, R. S., Tucker, D. M., Silberstein, R. B., & Cadusch, P. J. (1997). Eeg coherency: I: statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalography and clinical neurophysiology*, *103*(5), 499–515.

Oliveira, A. S., Schlink, B. R., Hairston, W. D., König, P., & Ferris, D. P. (2016). Proposing metrics for benchmarking novel eeg technologies towards real-world measurements. *Frontiers in human neuroscience*, *10*, 188.

OpenEEGProject. (2018). *Openeeg.* Retrieved from `http://openeeg.sourceforge.net/doc/` ([accessed: 13-09-2018])

Parra, L. C., Spence, C. D., Gerson, A. D., & Sajda, P. (2005). Recipes for the linear analysis of eeg. *Neuroimage*, *28*(2), 326–341.

Patel, S. H., & Azzam, P. N. (2005). Characterization of n200 and p300: selected studies of the event-related potential. *International journal of medical sciences*, *2*(4), 147.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology*, *49*(12), 1373–1379.

Shan, H., Xu, H., Zhu, S., & He, B. (2015). A novel channel selection method for optimal classification in different motor imagery bci paradigms. *Biomedical engineering online*, *14*(1), 93.

Shi, J. V., Xu, Y., & Baraniuk, R. G. (2014). Sparse bilinear logistic regression. *arXiv preprint arXiv:1404.4104*.

Stewart, A. X., Nuthmann, A., & Sanguinetti, G. (2014). Single-trial classification of eeg in a visual object task using ica and machine learning. *Journal of neuroscience methods*, *228*, 1–14.

Sur, S., & Sinha, V. (2009). Event-related potential: An overview. *Industrial psychiatry journal*, *18*(1), 70.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Tomioka, R., Aihara, K., & Müller, K.-R. (2007). Logistic regression for single trial eeg classification. In *Advances in neural information processing systems* (pp. 1377–1384).

Vallabhaneni, A., Wang, T., & He, B. (2005). Brain—computer interface. In *Neural engineering* (pp. 85–121). Springer.

Wang, Y., Zhang, Z., Li, Y., Gao, X., Gao, S., & Yang, F. (2004). Bci competition 2003-data set iv: an algorithm based on cssd and fda for classifying single-trial eeg. *IEEE Transactions on Biomedical Engineering*, *51*(6), 1081–1086.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical neurophysiology*, *113*(6), 767–791.

A Appendix

A.1 Consent Form

Teilnahme an der Studie: Traumschreiber als EEG-System

Hiermit bestätige ich, ___________________________ geboren am ______________________, dass ich dieses Dokument vollständig gelesen und verstanden habe. Mir ist bewusst, dass die Teilnahme an der Studie freiwillig ist und jederzeit ohne nachteilige Auswirkungen, oder Nennung von Gründen abgebrochen werden kann. Mir ist bewusst, dass die Speicherung der gemessenen Daten in anonymisierter Form erfolgt und bei möglichen Veröffentlichungen keine Rückschlüsse auf meine Person möglich sind. Ich wurde informiert, dass ich jederzeit meine Einwilligung zur Speicherung und Nutzung meiner Daten entziehen kann. In diesem Fall werden die Daten unverzüglich gelöscht.

Die gesammelten Daten werden folgenden Personen zur Verfügung gestellt: Merle Reimann, Ann-Kathrin Schalkamp und der Neuroinformatik-Gruppe der Universität Osnabrück.

Ich bin damit einverstanden, dass während des Experiments Messungen mit einem Elektroenzephalogramm (EEG) gemacht werden. Dabei werden vom Gehirn erzeugte Spannungen auf der Kopfoberfläche gemessen. Ich stimme zu, dass hierzu der Traumschreiber verwendet werden darf. Der Traumschreiber wurde von Johannes Leugering und Kristoffer Appel zur Erhebung von Daten während des Schlafs entwickelt. Diese Studie soll dazu dienen herauszufinden, ob der Traumschreiber auch als mobiles EEG-System verwendet werden kann.

Ich bestätige, dass ich keine neurologischen Erkrankungen habe.

Ich bestätige, dass ich die obige Erklärung vollständig gelesen und verstanden habe und ihr zustimme. Alle eventuellen Fragen wurden von den Versuchsleiterinnen beantwortet.

Datum, Unterschrift

aschalkamp@uos.de
mereimann@uos.de

A.2 AUC overviews

channel	0	1	2	3	5	6	7	8	dimensionality	training set
1	0.789892	0.760029	0.712143	0.722733	0.601468	0.760029	0.733229	0.743221	0.764063	0.766376
2	0.607308	0.650908	0.636707	0.583001	0.565509	0.650785	0.632993	0.596616	0.622468	0.639024
4	0.758368	0.756638	0.732226	0.682036	0.651401	0.731003	0.757018	0.743368	0.736067	0.730020
5	0.668265	0.687458	0.703304	0.713756	0.562684	0.715745	0.680428	0.650652	0.710241	0.736670
6	0.665390	0.715876	0.744021	0.576429	0.578668	0.679840	0.692760	0.681281	0.698255	0.701771
7	0.823851	0.821269	0.816025	0.661478	0.555633	0.804954	0.791178	0.806161	0.812362	0.820327
8	0.731557	0.727459	0.729508	0.631148	0.643443	0.754098	0.723361	0.713115	0.758197	0.761532
9	0.600929	0.627620	0.672044	0.568786	0.657647	0.580240	0.539241	0.576377	0.657344	0.639744
mean	0.705695	0.718407	0.718247	0.642421	0.602056	0.709587	0.693776	0.688849	0.719875	0.724433
1-8	0.725686	0.718576	0.731697	0.652273	0.544986	0.698674	0.706224	0.696262	0.731789	0.719447

Table A.1: SSVEP: AUC overview

channel	0	1	2	3	4	5	6	7	dimensionality	training set
1	0.537193	0.634208	0.660809	0.670077	0.686206	0.652744	0.685243	0.657198	0.653587	0.687290
2	0.557895	0.616667	0.614912	0.649123	0.632456	0.604825	0.739474	0.665351	0.655263	0.682357
4	0.451136	0.623214	0.640422	0.622240	0.632143	0.613312	0.746916	0.539773	0.539448	0.664487
5	0.586201	0.576299	0.550487	0.638636	0.694481	0.675812	0.639123	0.512825	0.576461	0.671514
6	0.542075	0.526457	0.588695	0.572844	0.573776	0.610956	0.647902	0.526807	0.495804	0.625337
7	0.620313	0.620313	0.677557	0.651562	0.594176	0.544886	0.663068	0.578551	0.605824	0.644341
8	0.535384	0.501984	0.592262	0.546296	0.583333	0.673280	0.516204	0.536045	0.552249	0.654211
9	0.506641	0.628874	0.641524	0.674889	0.668564	0.665718	0.575111	0.553289	0.631088	0.638939
mean	0.542105	0.591002	0.620833	0.628209	0.633142	0.630192	0.651630	0.571230	0.588715	0.658559
1-9	0.552278	0.586261	0.616140	0.658274	0.640484	0.649947	0.649477	0.547958	0.653318	0.667873

Table A.2: MP: AUC overview

channel	0	1	2	3	4	6	7	8	dimensionality	training set
1	0.852163	0.871394	0.700721	0.673077	0.712740	0.653846	0.735577	0.704327	0.680288	0.826738
2	0.771969	0.719132	0.685762	0.720245	0.515573	0.606229	0.706340	0.749166	0.641824	0.782883
4	0.634038	0.700779	0.754727	0.568409	0.536151	0.622358	0.669633	0.600667	0.586763	0.801200
5	0.631010	0.681490	0.554087	0.557692	0.546875	0.667067	0.697115	0.700721	0.618990	0.734826
6	0.687500	0.847222	0.854167	0.625000	0.465278	0.694444	0.680556	0.784722	0.694444	0.864311
7	0.571429	0.553571	0.642857	0.607143	0.535714	0.571429	0.464286	0.589286	0.535714	0.683233
8	0.639752	0.593944	0.758540	0.574534	0.643634	0.689441	0.653727	0.554348	0.701087	0.718814
9	0.634118	0.659412	0.623529	0.541765	0.639412	0.703529	0.694118	0.628824	0.644706	0.636264
mean	0.677747	0.703368	0.696799	0.608483	0.574422	0.651043	0.662669	0.664008	0.637977	0.756034
1-9	0.614154	0.586220	0.605031	0.589542	0.596217	0.602686	0.613897	0.640700	0.599025	0.653502

Table A.3: N240 target: AUC overview

channel	0	1	2	3	4	6	7	8	dimensionality	training set
1	0.852163	0.871394	0.700721	0.673077	0.712740	0.653846	0.735577	0.704327	0.680288	0.809343
2	0.771969	0.719132	0.685762	0.720245	0.515573	0.606229	0.706340	0.749166	0.641824	0.839450
4	0.634038	0.700779	0.754727	0.568409	0.536151	0.622358	0.669633	0.600667	0.586763	0.821209
5	0.631010	0.681490	0.554087	0.557692	0.546875	0.667067	0.697115	0.700721	0.618990	0.634782
6	0.687500	0.847222	0.854167	0.625000	0.465278	0.694444	0.680556	0.784722	0.694444	0.872555
7	0.571429	0.553571	0.642857	0.607143	0.535714	0.571429	0.464286	0.589286	0.535714	0.628670
8	0.639752	0.593944	0.758540	0.574534	0.643634	0.689441	0.653727	0.554348	0.701087	0.805475
9	0.634118	0.659412	0.623529	0.541765	0.639412	0.703529	0.694118	0.628824	0.644706	0.678424
mean	0.677747	0.703368	0.696799	0.608483	0.574422	0.651043	0.662669	0.664008	0.637977	0.761239
1-9	0.647439	0.621205	0.574887	0.580205	0.589590	0.575242	0.600705	0.632987	0.640599	0.639312

Table A.4: N240 distractor: AUC overview

Declaration of authorship

I hereby certify that the work presented here is, to the best of my knowledge and belief, original and the result of my own investigations, except as acknowledged, and has not been submitted, either in part or whole, for a degree at this or any other university.

Oelde, 26.09.2018 ....