Training Effects of a Tactile Brain-Computer Interface System During Prolonged Use by Healthy And Motor-Impaired People

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der
Julius-Maximilians-Universität Würzburg

Vorgelegt von
Matthias T. A. M. Eidel
geboren in Würzburg

Würzburg, 2020

Eingereicht am: 1. Mai 2020

<u>Mitglieder der Promotionskommission:</u>

Vorsitzender: Prof. Dr. Ricarda Scheiner

Gutachter: Prof. Dr. Charlotte Förster

Gutachter: Prof. Dr. Andrea Kübler

Tag des Promotionskolloquiums: 9. Juli 2020

Doktorurkunde ausgehändigt am: ....

Abstract

BACKGROUND - Brain-Computer Interfaces (BCI) enable their users to interact and communicate with the environment without requiring intact muscle control. To this end, brain activity is directly measured, digitized and interpreted by the computer. Thus, BCIs may be a valuable tool to assist severely or even completely paralysed patients. Many BCIs, however, rely on neurophysiological potentials evoked by visual stimulation, which can result in usability issues among patients with impaired vision or gaze control. Because of this, several non-visual BCI paradigms have been developed. Most notably, a recent study revealed promising results from a tactile BCI for wheelchair control. In this multi-session approach, healthy participants used the BCI to navigate a simulated wheelchair through a virtual apartment, which revealed not only that the BCI could be operated highly efficiently, but also that it could be trained over five sessions. The present thesis continues the research on this paradigm in order to

- confirm its previously reported high performance levels and trainability
- reveal the underlying factors responsible for observed performance increases
- establish its feasibility among potential impaired end-users

METHODS - To approach these goals, three studies were conducted with both healthy participants and patients with amyotrophic lateral sclerosis (ALS). Brain activity during BCI operation was recorded via electroencephalography (EEG) and interpreted using a machine learning-based linear classifier. Wheelchair navigation was executed according to the classification results and visualized on a monitor. For offline statistical analysis, neurophysiological features were extracted from EEG data. Subjective data on usability were collected from all participants. Two specialized experiments were conducted to identify factors for training.

Results and Discussion - Healthy participants: Results revealed positive effects of training on BCI performances and their underlying neurophysiological potentials. The paradigm was confirmed to be feasible and (for a non-visual BCI) highly efficient for most participants. However, some had to be excluded from analysis of the training effects because they could not achieve meaningful BCI control. Increased somatosensory sensitivity was identified as a possible mediator for training-related performance improvements. Participants with ALS: Out of seven patients with various stages of ALS, five could operate the BCI with accuracies significantly above chance level. Another ALS patient in a state of near-complete paralysis trained with the BCI for several months. Although no effects of training were observed, he was consistently able to operate the system above chance level. Subjective data regarding workload, satisfaction and other parameters were reported.

Significance - The tactile BCI was evaluated on the example of wheelchair control. In the future, it could help impaired patients to regain some lost mobility and self-sufficiency. Further, it has the potential to be adapted to other purposes, including communication. Once visual BCIs and other assistive technologies fail for patients with (progressive) motor impairments, vision-independent paradigms such as the tactile BCI may be among the last remaining alternatives to interact with the environment. The present thesis has strongly confirmed the general feasibility of the tactile paradigm for healthy participants and provides first clues about the underlying factors of training. More importantly, the BCI was established among potential end-users with ALS, providing essential external validity.

Zusammenfassung

HINTERGRUND - Brain-Computer Interfaces (BCI) ermöglichen ihren Benutzern die Interaktion und Kommunikation mit der Außenwelt, ohne dabei die Funktionstüchtigkeit der Muskeln voraus zu setzen. Zu diesem Zweck wird die Gehirnaktivität vom Computer direkt gemessen, digitalisiert und schließlich interpretiert. BCIs könnten daher eine wertvolle Methode sein, schwer körperlich beeinträchtigten oder sogar vollständig gelähmten Patienten zu assistieren. Viele BCI Ansätze basieren allerdings auf neurophysiologischen Potentialen, welche mittels visueller Stimulation evoziert werden. Dies kann zur Folge haben, dass das BCI von Patienten mit Sehbehinderung oder fehlender Kontrolle über die eigene Blickrichtung nicht erfolgreich benutzt werden kann. Deshalb wurden bereits einige nicht-visuelle BCI Paradigmen entwickelt. Insbesondere eine aktuelle Studie über ein taktiles BCI zur Rollstuhlkontrolle lieferte vielversprechende Ergebnisse: In fünf Trainingssitzungen navigierten gesunde Studienteilnehmer per BCI einen simulierten Rollstuhl durch eine virtuelle Wohnung. Hierbei konnte gezeigt werden, dass das BCI System nicht nur sehr effizient genutzt werden konnte, sondern auch, dass sich die Kontrolle durch das Training über mehrere Sitzungen verbesserte. Die vorliegende Dissertation befasst sich mit der weiterführenden Erforschung eben dieses Paradigmas, insbesondere mit den Zielen:

- die zuvor berichtete hohe Performanz und Trainierbarkeit zu bestätigen
- aufzuklären, welche Faktoren der Steigerung der BCI-Leistung zugrunde liegen
- die Anwendbarkeit des Paradigmas bei beeinträchtigten Endnutzern zu etablieren

METHODEN - Um diese Ziele zu erreichen wurden drei Studien sowohl mit gesunden als auch mit Teilnehmern mit amyotropher Lateralsklerose (ALS) durchgeführt. Während der BCI-Nutzung wurde die Gehirnaktivität per Elektroenzephalographie (EEG) aufgezeichnet und von einem linearen Klassifikator (basierend auf Maschinenlernverfahren) interpretiert. Die Navigation des Rollstuhls wurde entsprechend der

Ergebnisse des Klassifikators umgesetzt und auf einem Bildschirm visualisiert. Zur späteren statistischen Analyse wurden aus den EEG Daten neurophysiologische Merkmale extrahiert. Zudem wurden Fragebogendaten zur Nutzbarkeit des Systems von allen Teilnehmern erhoben. Zwei Experimente zur Identifizierung von Trainingsfaktoren wurden durchgeführt.

Ergebnisse und Diskussion - Gesunde Teilnehmer: Die Ergebnisse zeigten positive Effekte des Trainings auf die BCI Performanz und deren zugrundeliegenden neurophysiologischen Potentiale. Es konnte bestätigt werden, dass das Paradigma anwendbar und für die meisten Teilnehmer hocheffizient nutzbar war (im Vergleich zu anderen nicht-visuellen Ansätzen). Einige Teilnehmer mussten jedoch von der Analyse der Trainingseffekte ausgeschlossen werden, da sie keine ausreichende Kontrolle über das BCI ausüben konnten. Eine Steigerung der somatosensorischen Empfindlichkeitsschwelle wurde als ein möglicher Faktor für die Trainierbarkeit und Verbesserung der Performanz identifiziert. Teilnehmer mit ALS: Fünf von sieben Teilnehmern in verschiedenen ALS-Stadien konnten das BCI signifikant überzufällig benutzen. Ein weiterer ALS Patient mit nahezu vollständiger Lähmung trainierte den Umgang mit dem BCI über mehrere Monate hinweg. Er war beständig in der Lage, das System mit Genauigkeiten über dem Zufallsniveau zu steuern, jedoch konnten keine Trainingseffekte gezeigt werden. Fragebogendaten zur subjektiven Arbeitsbelastung, Zufriedenheit und einigen weiteren Parametern wurden ausführlich berichtet.

Bedeutung - Das taktile BCI wurde am Beispiel der Rollstuhlkontrolle evaluiert. In naher Zukunft könnte es beeinträchtigten Patienten helfen, ihre verlorene Mobilität und Selbstständigkeit zurück zu erlangen. Zudem kann es für viele weitere Zwecke adaptiert werden, insbesondere zur Kommunikation. Sobald visuelle BCIs oder andere technische Hilfsmittel bei Patienten mit (progressiver) motorischer Lähmung scheitern, könnten nicht-visuelle Paradigmen wie das taktile BCI zu den letzten verbleibenden Alternativen gehören, die eine Interaktion mit der Außenwelt noch erlauben. Die vorliegende Arbeit hat die grundsätzliche Anwendbarkeit des taktilen Paradigmas für gesunde Benutzer klar bestätigt. Zudem liefert sie erste Hinweise darauf, welche Faktoren den beobachteten Trainingseffekten zugrunde liegen könnten. Das BCI hat sich zudem bei potentiellen End-Nutzern mit ALS bewährt, was der externen Validität der Studienergebnisse enorm zuträgt.

Contents

List of Figures

List of Tables

1

Structure of This Thesis

This thesis revolves around the empirical research on a *Brain-Computer Interface* (BCI) paradigm intended for wheelchair control by severely paralysed people. The manuscript is structured in three major parts:

Part I - Theoretical Background provides a hierarchical introduction of all background information necessary to understand the research topic, starting out from a description of the group of potential end-users. This part will further introduce detailed information on BCI systems in general, and the paradigm of interest in particular, including how to conduct, measure and evaluate a BCI. Moreover, practical issues that may affect successful BCI operation, especially with regard to potential end-users, are outlined. Finally, the major research questions and goals are defined.

Part II - Empirical Investigation describes the three studies conducted to answer these questions. A visual abstract (Figure 1.1) provides an overview of this empirical investigation to illustrate the conceptual connection between the studies and their relevant research questions, as well as their principal results.

Each study will be reported with a short *introduction* and detailed *methods*, *results* and *discussion* sections.

Part III - General Discussion summarizes the major findings from all studies to derive a general conclusion and attempts to judge to which degree the research goals were achieved. Moreover, it outlines the limitations of the present BCI paradigm, but also provides important insights and recommendations for further applications. Finally, the thesis is concluded with an outlook on future studies.

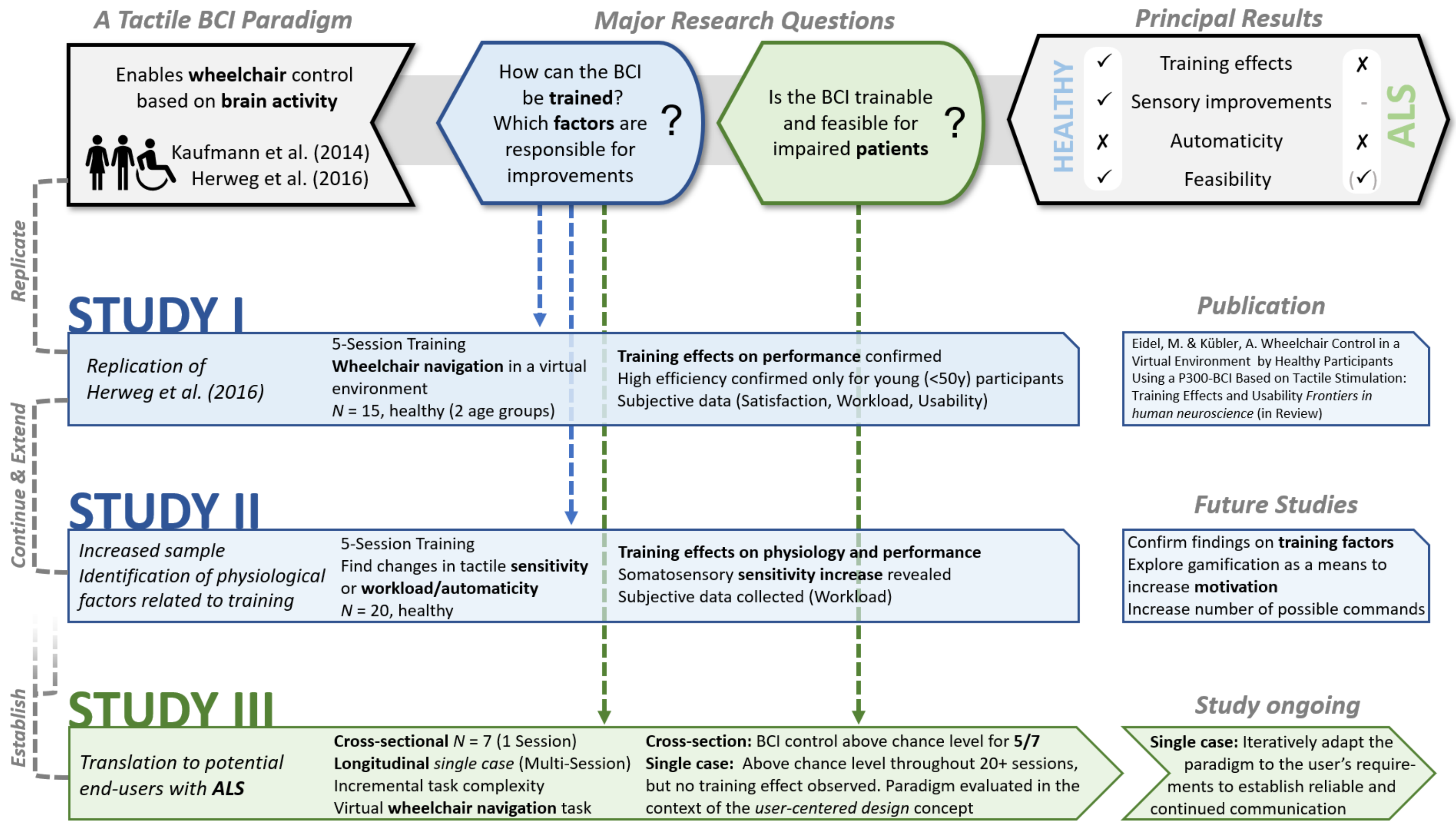

Figure 1.1: Visual abstract. BCI = Brain-Computer Interface. ALS = Amyotrophic Lateral Sclerosis.

Part I

Theoretical Background

2

The Locked-In Syndrome (LIS)

2.1 Definition and Causes

The state of a conscious mind within a paralysed body is broadly described with the term *Locked-In Syndrome* (LIS), which was first coined by Plum & Posner (1966). Their original definition described a state of near complete loss of motor function (tetraplegia and cranial nerve paralysis), with the exception of vertical eye movement, which was considered to remain intact. Some time later, Bauer et al. (1979) proposed a more refined definition of this neurological condition, and introduced three major subtypes of LIS according to their levels of motor impairment:

- Firstly, in the *incomplete locked-in syndrome*, some residual muscle movement persists, such as the ability to move a finger or simple facial expressions.
- Secondly, there is the *classical locked-in syndrome*, in which loss of voluntary motor function is nearly complete, leaving only vertical eye movements.
- Finally, there is the *total* or *complete locked-in syndrome (CLIS)*, where no voluntary muscle movements, including eyes and eyelids, remain possible.

Notably, consciousness is not considered to be affected by LIS or CLIS itself (Bauer et al., 1979). All the more, any degree of LIS is a severe medical disability that poses an array of challenges for the patients and their caretakers. Those include, but are not limited to, a need for assistance with communication and locomotion. Because of many superficial similarities, CLIS is especially difficult to differentiate from other causes of paralysis or disorders of consciousness, e.g. a vegetative state. In practice, it is often unclear what state of consciousness an unresponsive patient is in (Andrews et al., 1996). A technology to reliably discriminate these states and to assist those with intact consciousness would be most valuable.

There are a number of causes that can lead to a state of LIS or even CLIS - for in-depth information and reviews, see Laureys et al. (2005); Patterson & Grabois (1986). Briefly, among the most relevant diseases known to lead to LIS/CLIS are *spinal muscle atrophy*, and (brainstem) *stroke*. Other, less frequent causes include *multiple sclerosis*, *Parkinson's disease*, brain *hypoxia* (e.g. caused by heart attack) and other traumatic brain injuries. Overall, the multisystem disorder *Amyotrophic lateral sclerosis (ALS)* may be the most important cause of LIS/CLIS - especially in the context of the present work, since all included patients were suffering from some form of ALS at the time of data collection. Hence, ALS will now be introduced in more detail, and later chapters will be discussed with a focus on ALS. Nevertheless, results from this thesis may be transferable to similar medical conditions.

2.2 Amyotrophic Lateral Sclerosis (ALS)

Amyotrophic lateral sclerosis is a progressive, neuromuscular disorder in which upper and lower motor neurons degenerate, eventually leading to paralysis (for reviews see Kiernan et al., 2011; Mitchell & Borasio, 2007; Zarei et al., 2015).

There are two major subtypes of ALS, depending on where the first symptoms occur. In the "bulbar onset" form, patients usually exhibit motor-related difficulties with pronunciation of words (*Dysarthria*, which is unrelated to *Aphasia*, a problem with understanding or producing language) or difficulties with swallowing (*Dysphagia*). In the "spinal" or "limb onset" form, the first symptoms are often weakness, tremor or spasms in distal limbs such as arms and legs (Zarei et al., 2015).

Recent studies based on European ALS registers estimate the incidence of ALS at 2.6 cases, and ALS prevalence at roughly 8 cases per 100 000 people (Hardiman et al., 2017). Men are 1.5 times more likely to be affected by ALS and are also diagnosed at slightly earlier ages than women. ALS is diagnosed at median ages of 65.2 and 67 years for men and women, respectively. Incidence rates steadily increase after the age of 30 and decline after reaching a peak at 70-74 years for men and 65-69 years for women (Logroscino et al., 2010).

A study by Rooney et al. (2013) reported a median survival rate of ALS patients of 1.27 years from diagnosis (2.39 years from the onset of the first symptoms). Still, about 20 % of patients survive for 5 to 10 years after symptom onset (Talbot, 2009).

During the progression of ALS a complete paralysis of the body sets in, including respiratory and bulbar functions (Borasio et al., 1997; Zarei et al., 2015). This necessitates the decision for or against life-prolonging treatments such as artificial ventilation or artificial nutrition via percutaneous endoscopic gastrostomy (PEG). Regardless, even if these interventions are performed, ALS will continue to progress to more advanced stages, possibly leading to CLIS.

Despite years of research, the primary cause of ALS is not fully understood. Its occurrence is mostly sporadic and assumed to be linked to both genetic and environmental factors. It is assumed that only a minority of 5-10% of ALS cases are familial and inherited with a Mendelian pattern (Kiernan et al., 2011). While several gene mutations have been identified that can cause a clinical ALS phenotype, a conclusive understanding of its pathomechanisms is still lacking, perhaps because of the multitude of different factors and genetic pathways that may contribute to neural degeneration in ALS. Often, this degeneration is assumed to be caused by an activation of calcium-dependent enzymatic pathways induced by glutamate-induced excitoxicity (Heath & Shaw, 2002; Kiernan et al., 2011). Additionally, many other biomolecular contributers to ALS pathogenesis are described. Mitochondrial and ion channel dysfunction (Kanai et al., 2006; Sasaki & Iwata, 2007), cytoplasmic protein aggregation (Kepp, 2019) and the release of toxic or inflammatory molecules are among the factors most frequently found in literature (Kiernan et al., 2011).

Although the effects of ALS were once thought to be exclusive to motor neurons, it is now considered a multisystem disorder (Grossman, 2019; van der Graaff et al., 2009) that also affects sensory networks (Jamal et al., 1985). Furthermore, several studies have shown that there are effects on the central nervous system as well: For instance, the altered metabolism of glutamate described above seems to also affect the frontal cortex and spinal cord, as reported in several studies (Heath & Shaw, 2002; Plaitakis & Caroscio, 1987; Rothstein et al., 1992). Another effect on the central nervous system seems to be a reduction in glucose metabolism in the frontal lobe (Ludolph et al., 1992) and impairments of cognitive functions such as working memory and sustained attention (Hanagasi et al., 2002). Indeed, cognitive impairment, especially fronto-temporal dementia, is frequently associated with ALS (Kiernan et al., 2011; Mitchell & Borasio, 2007; van der Graaff et al., 2009). While this should be considered when working with ALS patients (especially when designing assistive paradigms for them), one should be careful not to generalize these findings, since the literature on cognitive impairments in ALS remains inconclusive. Several studies which have specifically addressed this topic could not demonstrate significant cognitive differences between late stage ALS patients and healthy controls (Lakerveld et al., 2008), or have found deficits in only 20-60 % of patients (Flaherty-Craig et al., 2006; Iversen et al., 2008).

A further non-motor related symptom of ALS concerns the structure of the skin. Here, certain alterations have been described in the context of ALS (Fullmer et al., 1960; Ono et al., 1990). These are likely due to a change in structure of collagen, a glycoprotein of the extra-cellular matrix of connective tissue, that is abundant in the dermis. In ALS, collagen fibrils have been found to have a smaller diameter and to be less densely packed as compared to healthy controls (Ono et al., 1990). The effects that this anomaly might have on tactile perception are still largely unknown.

Literature even provides evidence that the somatosensory neural pathways responsible for tactile perception can be affected in several places, but this will be explored later in chapter 4.1, which will provide an overview on human tactile perception. In any case, researchers should keep these effects of ALS in mind when designing or interpreting results from studies relying on tactile perception with ALS patients.

There is currently no cure for ALS, but there are some promising approaches that work on the aforementioned, glutamate-induced excitotoxicity to the nervous system. *Riluzole*, a glutamate-release inhibitor (Kanai et al., 2006), has been shown to extend life expectancy in controlled clinical trials, albeit only for a median time of three months (Miller et al., 2012). Because of this limited effect, its efficacy remains somewhat controversial and could not be demonstrated in several other studies (Dharmadasa & Kiernan, 2018). Nonetheless, it was approved by the US Food and Drug Administration in 1995.

Since this disease is not likely to be cured by pharmacological intervention in the near future, other approaches to assist the patients in their daily life are required. This is where *assistive technologies* (AT) such as BCIs become relevant. The present thesis will specifically explore a means to possibly assist with autonomous locomotion in the near future.

3

Principles of Brain-Computer Interfaces (BCI)

3.1 The BCI Cycle

BCIs allow for the human interaction with computers and thus the external world while relying only on brain activity (Kübler et al., 1999; Wolpaw et al., 1991). Ideally, neither intact muscle control nor verbal communication are required for controlling a BCI, rendering it a promising tool for patients with disabilities such as LIS. With the computer acting as an intermediary between brain and machine, almost anything can potentially be controlled via BCI: Applications range from simple *yes* or *no* responses over to full text spelling as a more refined means of communication, to the control of prosthetics, wheelchairs and other kinds of AT (Kübler, 2019; Vaughan, 2003).

The following sections will introduce key concepts and modes of operation of BCI technology. Specifically, they aim to provide the theoretical foundations that have been put into practice for the present work, starting from neurophysiological basics and recording of brain activity, to paradigm design and, finally, analysis and classification of the signal. The entire process from data acquisition to feedback and execution of commands is often called the *BCI cycle* (see Figure 3.1) (Gerven et al., 2009). For further reading, see Birbaumer & Cohen (2007); Kübler (2019); Kübler et al. (2001a); Wolpaw & Wolpaw (2012).

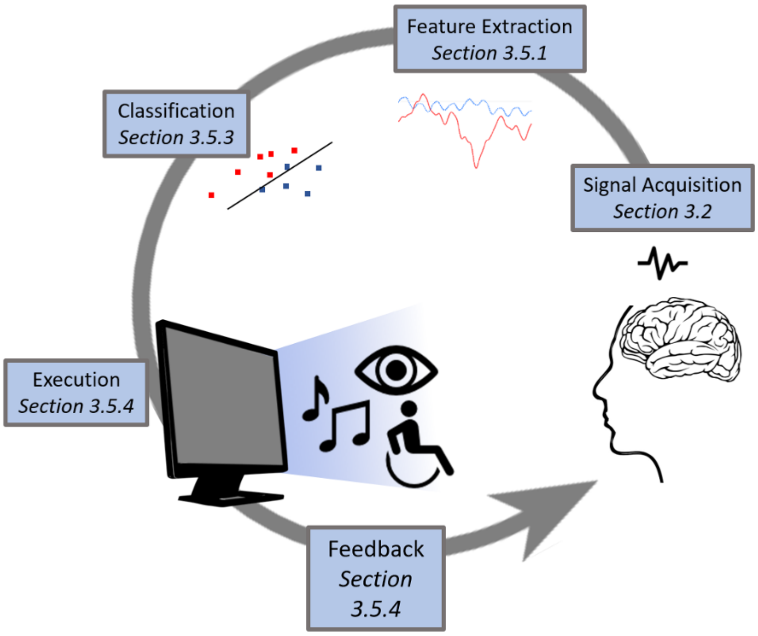

Figure 3.1: The BCI cycle from acquisition to execution and feedback. See the indicated sections for more details.

3.2 Recording Techniques: An Overview

In the context of BCI, brain activity can be measured with a variety of techniques:

- *Electroencephalography* (EEG)
- *Magnetoencephalography* (MEG)
- *Near-infrared spectroscopy* (NIRS)
- *Electrocorticography* (ECoG)
- *Intracortical microelectrodes* (IME)
- *Functional magnetic resonance imaging* (fMRI)
- *Positron-emission tomography* (PET)

All of these techniques have properties and advantages regarding portability, invasiveness and their respective temporal and spatial resolutions. These properties are broadly visualized in Figure 3.2.

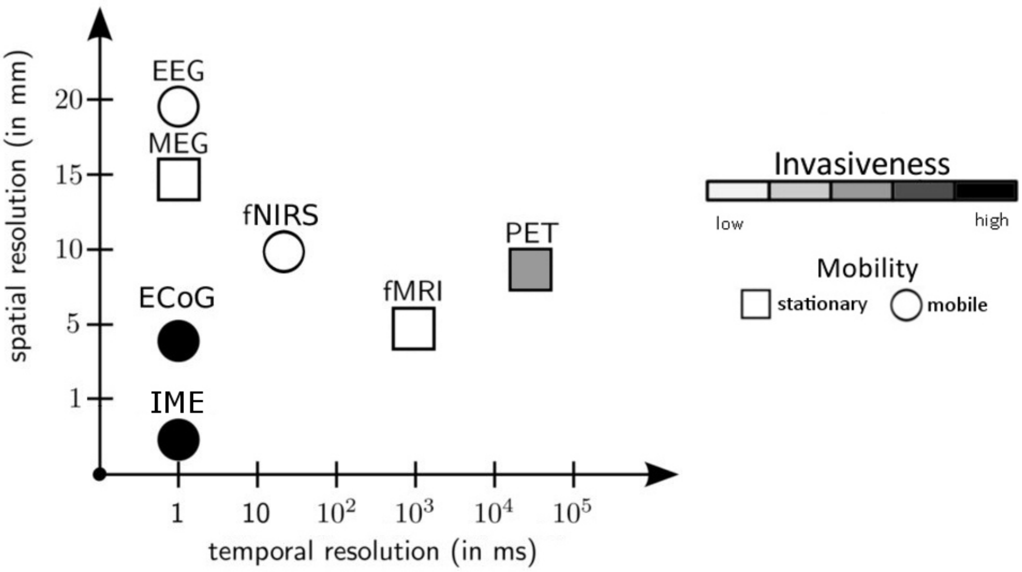

Figure 3.2: Selected techniques for measuring brain activity. Figure adapted from Hitziger (2015)

In the field of BCI, EEG has become the most abundantly used method because of its comparatively low cost, high transportability, high temporal resolution and non-invasiveness (Wolpaw & Wolpaw, 2012, p.105). Conversely, PET has found little application for BCI and will not be discussed further.

3.2.1 Electrical and Electromagnetic Recording

The following recording techniques rely on electrical activity in the brain, making them suitable for use cases in which a high temporal resolution is important. For an additional, comprehensive overview on these techniques, see Nunez et al. (2012).

EEG

EEG was discovered and first described by the German neurologist Hans Berger (1929). It is measured by placing one or several electrodes on the scalp surface. The so called *electroencephalogram* is then delineated by plotting voltage differences between a reference point and the measuring electrodes over time. Often, this reference is provided by a special electrode that is placed at a nearby, ideally inert (i.e. not affected by brain or muscle activity) position such as one or both of the *mastoids* or the tip of the nose. Alternatively, a *common average reference* can be used, in which the signal from all electrodes is mathematically averaged. Measurement electrodes are then referenced to this average (Luck, 2012).

In any case, detectable voltage differences are caused by the summed activity of large populations of cortical pyramidal neurons. Here, postsynaptic potentials of many synchronized neurons add up to dipols between apical dendrites and the soma. Those dipols can, if arranged perpendicular to the skull, be measured on the scalp surface. The responsible electrical signals propagate through tissue with comparatively small latencies, which accounts for the excellent temporal resolution of EEG. However, the spatial resolution of EEG is relatively low (in the range of few centimeters), since the signal blurs as it travels through a considerable distance of bone and tissue (Luck, 2012).

Electrode positions are internationally normed with the *10-20 system*, where positions are spaced 10 and 20 percent apart in respect of the front-back (*nasion to inion*) and left-right (*between the preauricular points*) dimensions of the skull (Jasper, 1958). All electrode positions described in the present work refer to this system.

ECoG and IME

Both *electrocorticography* (ECoG) and *intracortical microelectrodes* (IME) can record voltage from the same underlying physiological processes as EEG. Again, their temporal resolution is very high because of the electrical nature of the signal. Since electrodes or electrode arrays are placed directly onto or even into the cortex, spatial resolution of ECoG and IME is excellent. However, implanting such electrodes requires surgery which makes this a highly invasive technique that is reserved for a select few cases.

MEG

Signals of a similar origin as in EEG or ECoG can also be recorded with *magnetoencephalography* (MEG). Because of this similarity, temporal resolution is very high in MEG as well. Its spatial resolution is slightly superior to EEG, though still relatively low. Unlike for EEG however, the neurons measured have to be arranged tangentially to the cortical surface: An electromagnetic field emerges in a circular geometry around the direction of electric current that causes it. Therefore, activity from pyramidal cells in the *sulci* (i.e. the cortical folds), but not the *gyri*, can be detected outside of the skull. Here, MEG is measured with SQUIDs (super-conducting quantum interference devices) over the surface of the scalp. Because the MEG signal is orders of magnitude smaller (ranging between 10 fT to 1 pT) than the surrounding magnetic noise, the whole MEG-setup has to be placed in electromagnetically shielded rooms. These circumstances are what makes MEG both expensive and non-portable (Hämäläinen et al., 1993).

3.2.2 Hemodynamic Recording

Hemodynamic recording techniques rely on activity-dependent metabolic changes in the brain. These happen less immediately than electrical activity, as it takes about 2-6 seconds until a locally heightened metabolic demand is compensated with an increased cerebral blood flow which delivers oxygenated hemoglobin (Stippich, 2015). As a consequence of this metabolic latency, the respective hemodynamic techniques have a comparatively low temporal resolution.

Blood Oxygen Level-Dependent imaging (BOLD) was first described by Ogawa et al. (1990) and is now a common technique for identifying active brain regions. Brain regions that increase their level of oxygenated blood as described above can be measured and analysed in different ways. For a thorough layout of the following techniques, see Ramsey (2012); Stippich (2015).

fNIRS

The non-invasive *functional near infrared spectroscopy* (fNIRS) method uses near infrared light (690–830 nm), for which organic tissue is mostly transparent. For NIRS, several light emitters and detectors are placed on the head. Their light permeates the brain surface where it interacts with brain cortex tissue. For *functional* NIRS, hemoglobin and deoxygenated hemoglobin have different absorption spectra, so that when the refracted light enters the sensors, it is possible to calculate the relative hemoglobin/deoxy-hemoglobin levels which are indicative of brain activity. fNIRS is portable and can be used for BCI, despite its high latency (Sitaram et al., 2012).

fMRI

Functional magnetic resonance imaging (fMRI) is another technique that can measure and visualize brain activity using the BOLD method in great detail. As the name suggests, MRI is an *imaging* technology that relies on principles of magnetic resonance (for the detailed physics of its functionality, see McRobbie et al., 2017). While MRI can produce high resolution images of brain anatomy, *functional* MRI integrates BOLD contrasts to infer brain activity. Here, levels of hemoglobin/deoxy-hemoglobin can be measured quantitatively making use of their different magnetic properties.

Although fMRI has become much faster due to the increase of computational power and the development of more sophisticated algorithms, its temporal resolution remains as low as limited by the metabolical response of the subject. With the exception of a few special applications that require the use of contrast agents, MRI is completely non-invasive.

3.3 Subtypes of EEG based BCI

Since the BCI studies in the present work exclusively use EEG, the following descriptions will focus on some of the most important BCI approaches based on this technique.

There is a variety of possible classifications of EEG paradigms. One important distinction can be made between signal types of *active* BCIs that can be used proactively by the user without the need of external stimulation, and *reactive* BCIs, which do depend on external stimuli to which the user reacts. One example for the former is given by interfaces based on *sensory motor rhythms* (SMR):

Active

SMR BCI takes advantage of the fact that, over sensorimotor cortices of the brain, the EEG signal in the alpha (8-12 Hz) and beta (18-30 Hz) frequency bands can be voluntarily modulated. Chatrian et al. (1959) discovered an (event-related) *desynchronization* of neurons in a motor area and a consequent decrease of the mu rhythm during movement. Depending on motor activity and motor area of interest, SMR can also *synchronize* (i.e. increase in band power), an effect that has been associated with 'idling' or inhibition (Pfurtscheller, 1992; Pfurtscheller & Da Silva, 1999). SMR modulation can be achieved solely by imagining a movement (Neuper & Pfurtscheller, 1999), making it useful for controlling prosthetic limbs (Neuper et al., 2006) and BCI in general. SMR based BCI requires practice, but it has been shown to work even with severely handicapped ALS patients (Kübler et al., 2005).

Reactive

Examples for *reactive BCI*, on the other hand, include *steady-state evoked potentials* (SSEP) and *event-related potentials* (ERP).

The most prevalent SSEP BCI applications are based on steady-state *visually* evoked potentials (SSVEP; for an overview, see Faller, 2012). Their underlying principle is that, when visually stimulating with a stable oscillation of a fixed frequency, this same frequency (and its harmonics) can be found as voltage oscillations over associated brain areas - in this case, the occipital cortex. BCI paradigms which present several visual stimuli at the same time, but with distinct frequencies, can make use of this effect: A study by Faller et al. (2010), for instance, uses frequency analysis to determine which, if any, stimulus is observed by the BCI user, thus enabling them to make selections. There are also auditory (SSAEP) (Hill & Schölkopf, 2012) and, notably, tactile or somatosensory (SSSEP) (Ahn et al., 2016; Erp & Brouwer, 2014; Pokorny et al., 2016) variants and BCI applications of steady-state potentials.

The following studies are entirely based on ERP detection. An ERP is a reliable brain response that is elicited by an event, usually an external stimulus. ERPs can easily be measured with EEG and similar techniques. Although ERPs had been recorded before, the scientific interest was truly sparked in 1964, when a cognitive ERP component, the *contingent negative variation*, was first described (Walter, 1964). Only shortly thereafter, Sutton et al. (1965) published their discovery of the P300 ERP, which has become one of the most important and intensely researched ERPs (Luck, 2005, p. 5). Therefore, we will now take a closer look at the P300.

3.4 The Oddball Paradigm and the P300 ERP

ERPs are often named according to their polarity and latency. Thus, this ERP, appearing about 300 ms after the onset of a stimulus as a positive deflection, was named *P300*. The P300 can be elicited with visual stimuli, but also with auditory and tactile stimuli if a visual paradigm is not feasible (see section 5.2.1 for the gaze dependency of visual P300).

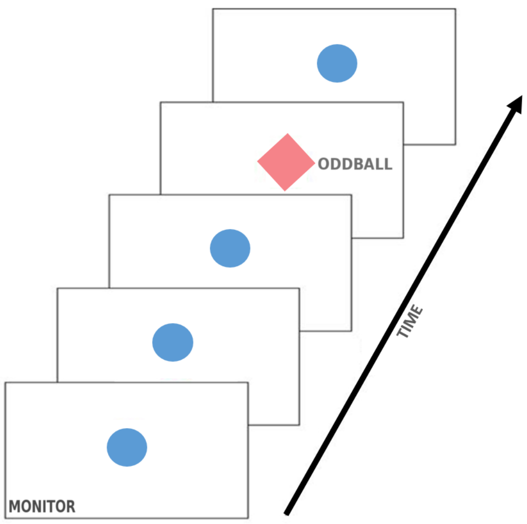

Figure 3.3: In the oddball paradigm, the target stimulus appears randomly within a sequence of frequent non-targets (adapted from Daltrozzo & Conway, 2014).

There are specific conditions under which the P300 can be observed: There should be at least two categories of distinguishable stimuli, the rare *target* (sometimes also known as deviant) and the frequent *non-target* (sometimes also known as standard) class. Targets should occur randomly within a sequence of non-targets with a probability of, ideally, 20 % (Cohen & Polich, 1997), although lower probabilities have been shown to result in larger ERP amplitudes. Since the P300 is strongly modulated by attention, the participant should focus on the occurrence of the target stimuli while ignoring the non-targets (Wickens et al., 1980). If attention is not appropriately allocated to the targets, the P300 amplitude will be decreased (Ford et al., 1976; Rappaport et al., 1990). This can be achieved experimentally by using a dual-task paradigm, in which participants have to allocate attention to a second, unrelated task (Kida et al., 2004; Käthner et al., 2014).

A prevalent P300 design that implements these principles is known as the *oddball paradigm* (Donchin et al., 1978, p. 349-411), a simple discrimination task where the participant is instructed to pay close attention or specifically respond to a randomly occurring, rare target stimulus, the eponymous "oddball", within a sequence of frequent non-target stimuli (Figure 3.3). One prevalent strategy to facilitate concentration on the target is to instruct participants to mentally count along with or to press a button at the occurrence of the target stimulus. This paradigm has been extensively used for neurophysiological research and is considered to elicit P300 deflections very reliably (Huettel & McCarthy, 2004). Further, we know that it can also be used with motor-impaired patients (Kleih et al., 2011).

Figure 3.4 depicts one example of a visual P300 deflection. Anatomically, the P300 has a centro-parietal distribution and is measured with the highest amplitudes at electrode positions Fz, Cz, Pz (Duncan et al., 2009). It is known that P300 amplitudes decrease as a result of a reduced ability to discriminate stimuli (Comerchero & Polich, 1999), which is important considering that this is likely the case for some ALS patients (see 4.1). Hence, stimuli should be well perceivable and easy to discriminate.

The P300 can be divided in several subcomponents, most commonly the P3a and P3b. P3a is sometimes dubbed the "novelty" P300 and is characterized by a shorter peak latency and a distribution shift toward central/parietal areas. P3b is described as a ”task-relevant” P300 component and strongest over parietal areas (Comerchero & Polich, 1999; Polich, 2007).

When used in a BCI paradigm, many P300 epochs should be averaged to increase the signal/noise ratio of the data. Studies reported that as few as 36 (Duncan et al., 2009) or even 20 epochs (Cohen & Polich, 1997) sufficed to yield statistically stable P300 deflections, though in practice, a higher number is recommended.

Based on the P300 and its fundamental principles of elicitation derived from the oddball paradigm, more elaborate designs can be created and used for BCI. A popular implementation is the visual spelling matrix established by Farwell & Donchin (1988), but there are many other paradigms, both visual and non-visual, that are based on the very same principles. These examples will be explored separately in chapter 5.

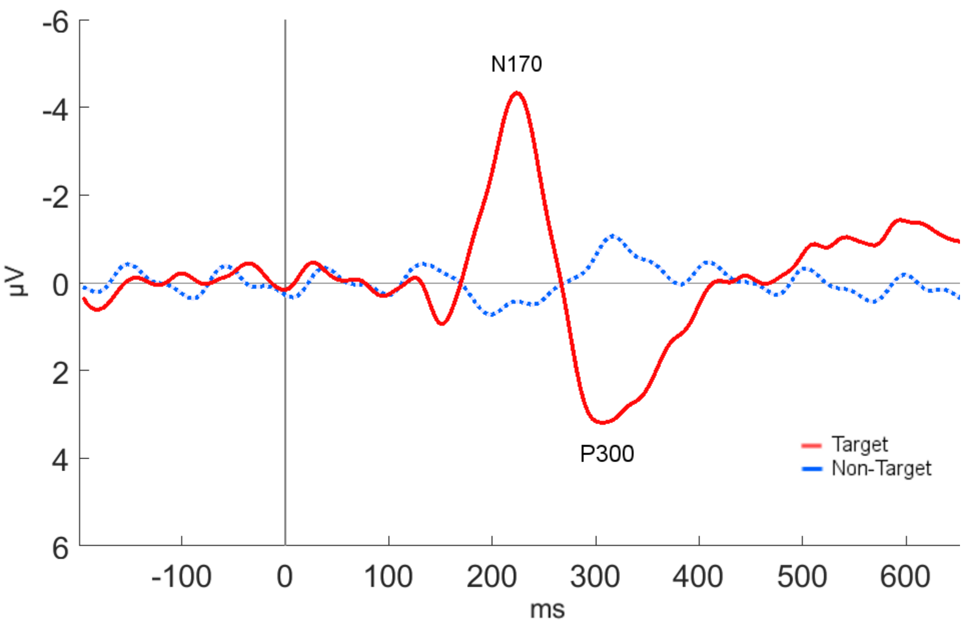

Figure 3.4: Exemplary illustration of a visually evoked P300 deflection. Also visible is the N170 ERP (see section 5.1 for details).

3.5 BCI Analysis

3.5.1 Measures for ERP Evaluation

For statistical analysis and ERP classification via machine learning, a number of primary variables can be derived from the EEG signal. Depending on the research question, these can be calculated either from individual or grand averages (i.e. over all included participants), but also on a single-ERP basis. These measures are:

- **Mean amplitude** of the ERP deflection in a defined post-stimulus time window. This window is chosen depending on the ERP component of interest. In the example shown in Figure 3.4, a window of 250-400 ms could be chosen to analyse the P300. The average of all samples within the window is extracted as mean amplitude.
- **Mean difference** of target and non-target epochs in a relevant time window as a measure of divergence. First, the difference curve is calculated by subtracting the non-target signal from the target signal. Then, the average of all samples within the window is extracted.
- **Peak amplitude**, which is defined as the highest or lowest signal value (depending on the ERP). This measure is discouraged by some researchers because of its lack of robustness. Instead, mean amplitude in a relevant time window around the peak should be used (Luck, 2005, p. 229-235).
- **Peak latency** is the time between stimulus onset and the peak. It is often detected automatically in a given time window and also potentially prone to a lack of robustness.

3.5.2 Measures of BCI Performance

Furthermore, a number of parameters can be used to describe effectiveness and efficiency of BCI control. These are derived from the primary measures using classification algorithms:

- **Online accuracy**, defined as the fraction or percentage of correct classifications during online use of the BCI. For effective BCI control, accuracy should be at least 70 % (Kübler et al., 2001b).

- **Offline accuracy** is the accuracy as determined after online BCI use, sometimes using extensively processed data and optimized classifiers. Offline accuracy should be cross-validated to pre-empt overfitting.

- **Information transfer rate** (ITR) is given in *bits per minute*. In addition to accuracy, it also takes into account the time required as well as the number of possible selections per trial. A high value indicates accurate and fast (hence, efficient) BCI control. For the mathematical definition, see equation 3.1. The formula assumes equal probability of all possible selections.

$$Bits = log_2(N) + A \cdot log_2(A) + (1 - A) \cdot log_2\left(\frac{1 - A}{N - 1}\right)$$
$$ITR = Bits \cdot SpM \tag{3.1}$$

N = number of possible selections; A = accuracy; SpM = selections per minute

- **Utility rate**, as a more user-centred alternative to the ITR (which focuses on classification, but neglects how well the user interface is designed). The utility rate takes both the classifier *and* characteristics of the interface into account - for instance, how error correction is handled - to provide a more practical metric of the system's usefulness (Seno et al., 2010).

3.5.3 Classification via Machine Learning

To allow for immediate feedback and execution of a command, specialized *machine learning* algorithms are employed during online BCI use. These aim to decode the user's intention based on features extracted from the raw data.

In the specific case of ERP BCI, a classification algorithm has to discriminate post-stimulus epochs between two classes: target or non-target. One widespread approach to do this is classification via *support vector machines* (SVM) or *linear discriminant analysis* (LDA).

The software used in the present work, BCI2000 (Schalk et al., 2004), implements a variant of the LDA algorithm, the *stepwise* LDA (swLDA) (Schalk & Mellinger, 2010, p. 208-213). Here, *feature weights* are calculated by stochastically introducing and removing spatiotemporal features from target and non-target epochs (features being specific EEG channels and time windows) into the calculation of a linear discriminant function based on multiple regression (Krusienski et al., 2008). As implemented in BCI2000, this iterative process attempts to optimize the model by trying to maximize the R^2 value (the coefficient of determination). Technically, this value describes proportion of variance in the data that is explained by a model. In simplified terms, the R^2 can be thought of as a measure of how well the classes can be distinguished by the model, with high R^2 values indicating that the underlying features allow for a reliable classification. As visualized in Figure 3.5, a strong divergence of target and non-target curves is usually mirrored in a high discriminability of the classes. It is important to understand that this value does not simply derive from the difference of the averaged curves (which might be strongly affected by noise and outliers). Instead, a small but robust divergence can result in higher R^2 values than a larger divergence observed in noisy data.

The process of defining feature weights on annotated data is known as *calibration.* Typically, several calibration runs are performed to account for user-specific characteristics in the EEG response before the system is ready to achieve reliable classification. During these runs, the user is instructed to concentrate on specific targets, so that their identity can be annotated in the EEG data for the subsequent classifier training. The thus calculated discriminant function can then be used to classify responses during online BCI, without prior knowledge of the class identities (Krusienski et al., 2006).

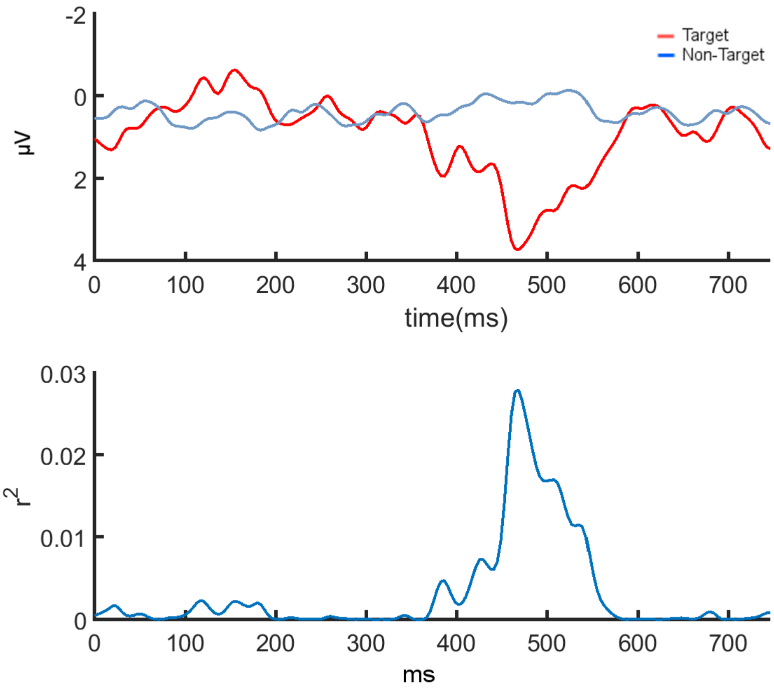

Figure 3.5: Example of the R^2-value during post-stimulus epochs. The divergence of curves between 400 and 600 ms is reflected in a high discriminability of the classes.

3.5.4 Feedback

Finally, the computer supplies feedback about the results of the current classification, thus concluding the BCI cycle. Swift feedback is not only necessary for facilitating the learning process, memorization and motivation (Anderson, 2000) - in BCI, the feedback is often identical to the very purpose of the system, meaning that it is given in form of the letter, the movement, or the command that was selected by the computer's classification. This feedback of the classification results enables the user, among other things, to identify and correct wrong or otherwise unintended commands. It can be provided visually or auditorily, or even over the sense of touch and a combination of these modalities.

4

Considerations for BCI Application in Practice

Having introduced these general aspects of BCI, this chapter will now widen the focus and explore further aspects that are no less crucial for understanding, evaluating and optimizing a BCI paradigm.

As the present work specializes on BCI paradigms using vibrotactile stimuli, it is important to understand how these stimuli are processed. Thus, the next section will summarize human somatosensory perception.

4.1 Somatosensory Perception

The biophysiological basics of somatosensory perception are exceptionally relevant in the context of tactile BCI paradigms. This section will give a brief overview of the various mechanoreceptors found in human skin (Table 4.1) and their possible relevance within the tactile paradigm. Known effects of age will be outlined, as ALS patients are often among the elderly (ALS incidence peaks at about 70 years of age - see section 2.2).

Furthermore, a well-established human somatosensory pathway (Figure 4.1), as described in Kandel's *Principles of Neural Science* (Kandel et al., 2000, Chapters 22-23) will be summarized. Evidence for effects of ALS on the receptors and pathways will be discussed in parallel.

4.1.1 Human Mechanoreceptors

Table 4.1: Overview of relevant mechanoreceptors.

Receptor	Location	Responds to
Merkel Discs	hairy and hairless skin	slow, punctate pressure
Ruffini Corpuscles	joints & skin	pressure and stretching
Meissner Corpuscles	hairless skin	pressure change (10-50 Hz)
Pacinian Corpuscles	concentrated on hands and feet	vibration (20-1000 Hz)

One type of receptors thought to contribute to human tactile perception are the **Merkel Discs**, which are located in the basal layer of the epidermis of both hairy and hairless skin (Halata et al., 2003). These receptors are connected to myelinated nerve fibers via a structure that is assumed to work like a synapse. Merkel Discs are slowly adapting mechanoreceptors which react mostly to punctate pressure, but also to the bending of hair (in some mammals, Merkel receptors are especially abundant at the base of the whiskers). Since the tactile paradigm of the present work uses medium to high frequency vibrations, the Merkel Discs, being not sensitive in this frequency range, may not be particularly important.

Another receptor type, the spindle-shaped **Ruffini Corpuscles** or sometimes **Bulbous Corpuscles**, are found in the deeper (reticular) layers of the skin. Notably, these spindles are also found in joint capsules. In either case, they respond to sustained pressure and stretching, possibly contributing to proprioception (Halata, 1988; Halata & Munger, 1981). Again, sustained deformation or proprioception are likely not especially relevant in the context of vibrotactile paradigm used in this work.

Meissner Corpuscles on the other hand, are specialized encapsulated nerve endings that react to both touch and low frequency vibrations (10-50 Hz) (Piccinin & Schwartz, 2018). They are located in the dermal papillae (which are small extensions into the epidermis) of hairless skin. Since they are highly sensitive to even minute pressure changes, they are often associated with the human ability to read Braille. Meissner Corpuscles have been found to decline with age (Iwasaki et al., 2003), a fact that should be considered when working with elderly end-users of tactile BCI systems, who may experience a reduced tactile sensitivity.

Finally, the **Pacinian Corpuscles** are primary mechanoreceptors, meaning that there is no synapse between receptor membrane and the locus of action potential generation (Bell et al., 1994; Davis, 1961). They are most abundant in the subcutaneous and dermis layers of the hairless skin of hands and feet, where they are sensitive to mechanical deformation and vibrations between frequencies of 20-1000 Hz. Hence, they play a major part for the sense of touch - in fact, they have been described to be up to 30 times more sensitive than the non-Pacinian receptors above (Bell et al., 1994; Bolanowski Jr & Zwislocki, 1984). Again, the number of Pacinian Corpuscles is known to decline with age (Cauna, 1964), with the same implication of a possibly reduced sensitivity among the elderly. Still, their sensitivity to vibrotactile stimuli with the frequency used in the tactile paradigm (250 Hz, see section 7.1 for details) may render them particularly relevant.

Overall, little is known about whether ALS directly affects mechanoreceptors or about the relevance of the changed collagen structure in the skin (which was outlined in section 2.2). Advanced age, however, seems to have a negative influence on the receptors (for a comprehensive review, see Wickremaratchi & Llewelyn, 2006). Further, there is evidence that the decline of the receptors indeed manifests in reduced sensory sensitivity. For instance, a study by Gescheider et al. (1994) found that thresholds for detecting vibrotactile stimulation increased with age. More specifically, a recent study by Chen et al. (2019) compared EEG responses to vibrotactile stimuli between age groups, and reported a power reduction in the EEG measures and lower BCI accuracies of the elderly. Clearly, tactile perception begins at the receptors, but it does not end there. Thus, the sensory pathway that is most relevant for the transmission of tactile sensation to the brain will now be introduced.

4.1.2 The Dorsal Column-Medial Lemniscal System

Signals from the receptors are relayed to the thalamus and cerebral cortex via somatosensory pathways. This section will outline the *dorsal column-medial lemniscal system* (see Figure 4.1), a major ascending pathway which is functionally specialized for the transmission of tactile and proprioceptive information (Kandel et al., 2000).

Once a mechanoreceptor is activated by an appropriate stimulus on the skin, its sensory information is then encoded by bipolar neurons of the **dorsal root ganglia**. These primary afferent fibers innervate only one type of receptor cell. Usually, one neuron bundles several similar receptors into small receptive fields. Likewise, multiple neurons form a nerve that bundles many fibers from larger areas of the skin called the *dermatomes*. Despite this aggregation, any functional separation of the afferents is conserved until they reach higher centers of the brain.

In this and similar sensory pathways, neurons that encode different sensory modalities (i.e. touch versus temperature or pain) differ in regard of their transmission speeds, which again are dependent on the axons' myelinisation and diameter. For instance, large diameter Aα and medium diameter Aβ fibers (both myelinated) are known to convey signals from each of the mechanoreceptors introduced above.

Negative effects of ALS may appear even at this stage of early somatosensory nerve transmission:

- A meta analysis by Hammad et al. (2007) found pathologic sensory abnormalities, e.g. axonal degeneration, in 91 % of ALS patients. Within those, mostly large-diameter axons (73 %), but also smaller axons were affected (23 %).
- There is strong evidence that non-motor nerve conduction is impaired in ALS patients. Compared to healthy controls, up to 72 % of patients showed abnormal nerve conduction in a study by Isak et al. (2016). The authors linked this to an involvement or even axonopathy of distal Aβ sensory fibers.

Next, a segregation of functionally distinct nerves occurs just as they exit the dorsal root ganglion and enter the **spinal column** through ipsilateral nerve openings in the vertebrae: Some branches of Aα fibers terminate in the ventral horn, whereas Aβ fibers terminate in different layers of the dorsal horn of the spinal cord. Of those, layers III to V contain secondary neurons that further relay the signal to the brain stem and the thalamus.

There is some evidence that ALS also disturbs the pathway at the spinal level:

- High levels of oxidative damage, indicative of neurodegeneration, were identified in spinal cord samples of ALS patients, specifically in the anterior and dorsal horns (Niebrój-Dobosz et al., 2004). Further, a histological study by Williams

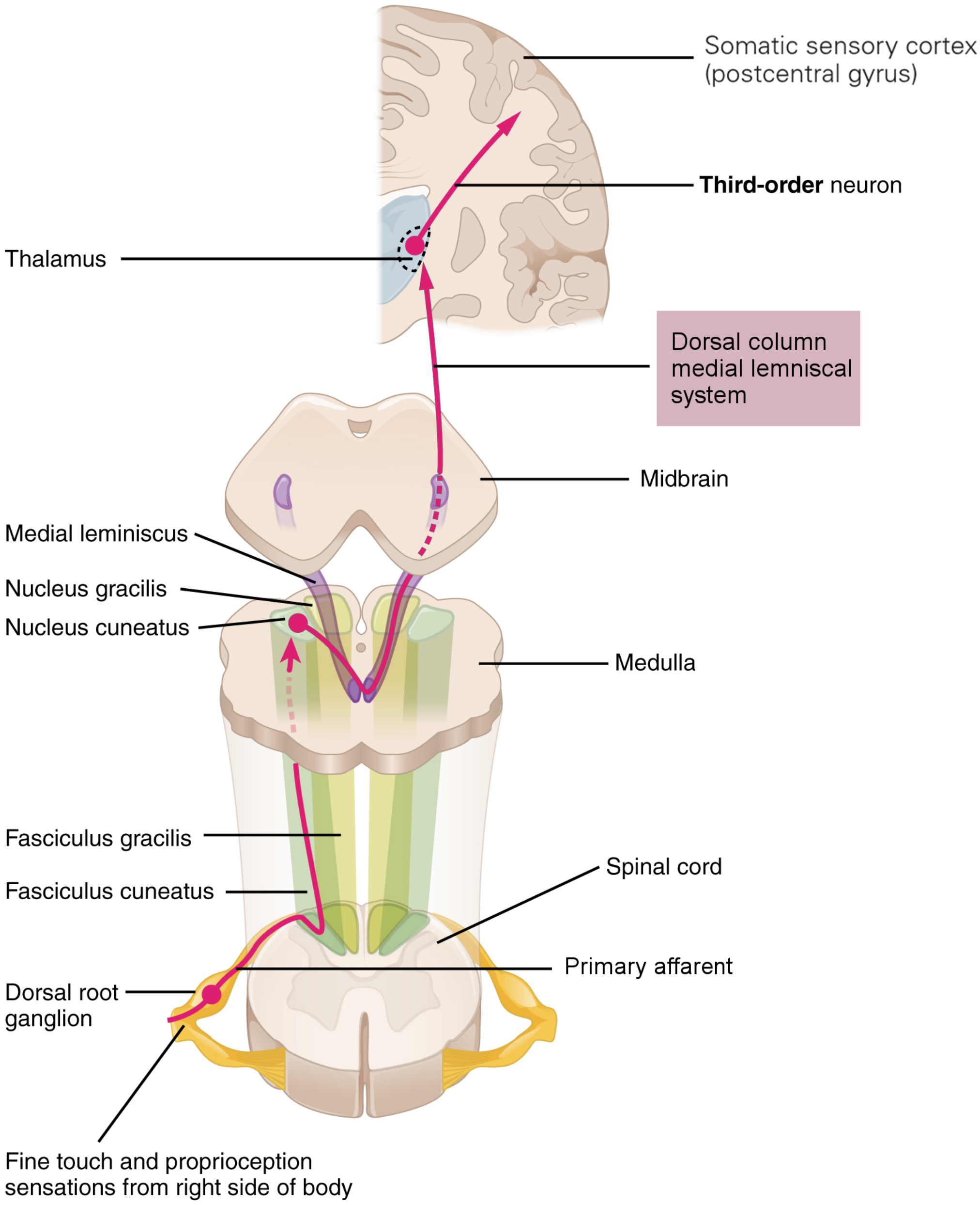

Figure 4.1: Overview of the dorsal column-medial lemniscal system, a major ascending somatosensory pathway (Betts et al., 2013). Primary afferents enter and ascend the spinal column, cross the midline in the medulla and enter the thalamus, which further projects to the primary somatosensory cortex.

The figure was modified from OpenStax College, provided under the terms of the Creative Commons Attribution 3.0 Unported License. Access for free at https://openstax.org/books/anatomy-and-physiology/pages/1-introduction.

et al. (1990) provided evidence that the spinocerebellar pathway is affected in sporadic ALS. Neither study specifically addressed the sensory system, but at the spinal level, the spinocerebellar and dorsal column-medial lemniscal pathways are anatomically similar.

- Cohen-Adad et al. (2013) found significant electrophysiological and pathological evidence for impairment and atrophy of sensory fibers in the dorsal column in the spinal cord of early-stage ALS patients. Again, large-caliber myelinated fibers were affected the most.

Continuing on the pathway, both Aα and Aβ fibers in the spinal cord also send a major ascending branch through the dorsal columns, more specifically through the *fasciculus cuneatus*, to the medulla (Figure 4.1). Here, second-order fibers join into the *medial lemniscus*, an ascending bundle of myelinated axons. This structure crosses the medulla's midline, so that from this point on, the pathway continues contralaterally. This crossing of somatosensory fibers is also referred to as *sensory decussation*.

After that, the fibers transverse the midbrain and finally terminate in the **thalamus**. Here, somatosensory information from the skin is integrated in the lateral and medial ventral posterior nuclei. Even now, a distinct somatotopic order of input is retained, so that individual body parts are represented by specific clusters of neurons in these nuclei. Finally, these nuclei of the thalamus convey tactile (and proprioceptive) information toward the different representative areas of the **primary somatic sensory cortex**. With this, the ascending somatosensory pathway is concluded.

- Somatosensory evoked potentials (SEP) can be measured as a cortical EEG response to stimulation. Since they are generated toward the end of the sensory pathway, they are often used to assess sensory transduction abnormalities (Iglesias et al., 2015), which could arise at any point upstream of the pathway.

- Iglesias et al. (2015) also found that in ALS, alterations (reduced amplitude or latency delay) of the SEPs were significantly correlated with disease duration.

- Overall, EEG experiments showed abnormal SEPs in 60 % - 72 % of ALS patients (Bosch et al., 1985; Iglesias et al., 2015). These impairments could not be attributed to peripheral conduction, which was not significantly affected in either study.

- Rather, there is ample evidence (both pathological and MRI based) that thalamocortical projections are frequently degenerated in ALS (Bede et al., 2013; Bosch et al., 1985; Brownell et al., 1970; Smith, 1960).
- Georgesco et al. (1997) and Sangari et al. (2018) showed again that SEPs were severely impaired in ALS. They suggested this was due to abnormal cortical integration involving thalamocortical projections and possibly pyramidal system dysfunction and cortical excitability changes. However, an effect on SEPs could also be a result from abnormalities of the dorsal nuclei of the spinal cord, from much earlier in the pathway (Georgesco et al., 1997).

In summary, it seems abundantly clear that ALS is not restricted to motor neurons. This short overview of the dorsal column-medial lemniscal system demonstrates that ALS can have a negative effect at many different levels of a somatosensory pathway, which is something that should be kept in mind when interpreting results from a tactile BCI used by ALS patients. Ideally though, BCI paradigms should already be designed with stimuli that are easily perceivable even with sensory impairments, so that minor sensory impairments may be compensated.

4.2 The User-Centred Design (UCD)

BCI paradigms and the P300 itself have been extensively researched - though mostly with healthy volunteers in the laboratory (Allison & Neuper, 2010). A majority of these studies, however, put emphasis on optimizing classification algorithms and other technical facets, or on mere proofs of concept, but fail to adopt a user-centred perspective (Pasqualotto et al., 2012).

While basic research is a cornerstone of science and fundamental to all subsequent applications, it is important to not lose focus on the intended end-users of the technology. Scientific studies on LIS/CLIS individuals are rather scarce, mainly because of the limited accessibility of the patients (Kübler et al., 2014). Unfortunately, results from a healthy population do not necessarily translate into the use with patients with ALS or other severe disabilities. Patients may have completely different, individual requirements. In fact, there is a translational gap between the laboratories and the patients in need, who rarely get the chance to use BCI in their daily life (Kübler, 2013; Kübler et al., 2015; Nijboer, 2015).

4.2.1 Applying UCD to BCI

To alleviate this issue, some researchers have included potential end-users in long-term studies, for instance Holz et al. (2015a,b); Sellers et al. (2010), who demonstrated the feasibility of prolonged and independent home use of BCI.

It is assumed that there is no "universal BCI" system that works for everybody (Allison & Neuper, 2010). Hence, another important approach to bridge the translational gap is to systematically redesign the BCI development and research process to acknowledge the user's biopsychological and social requirements by applying a *user-centred design* (UCD) concept - as defined in ISO-9241-210 (2010). This standard defines six important principles of UCD (quoted verbatim):

- the design is based upon an explicit understanding of users, tasks and environments
- users are involved throughout design and development
- the design is driven and refined by user-centred evaluation
- the process is iterative
- the design addresses the whole user experience
- the design team includes multidisciplinary skills and perspectives

After the first BCI study that implemented this approach with patients was published in 2011 (Zickler et al., 2011), it gradually gained traction within the BCI community. There have been calls to further establish UCD so as to maximize BCI usability (Nijboer, 2015), and to include the target group into the process, who might otherwise abandon the technology (Kübler et al., 2015).

Kübler et al. (2014) specifically adapted the UCD concept for BCI and applied the approach to 19 patients with severe motor impairments. In this adaptation, physiological and performance measures (as introduced in section 3.5.2) were included and extended with questionnaires to operationalize all aspects of usability: effectiveness (accuracy and NASA Task Load Index (TLX)), efficiency (ITR and workload) and satisfaction (rating scales and questionnaires).

Several questionnaires were suggested for use in UCD adherent BCI workflows to assess usability, workload and user experience, thus enabling researchers to iteratively tailor the paradigms to the needs of the individual end-user (Kübler et al., 2014).

The next section will establish the theoretical background of these questionnaires, while their specific application will be outlined in the methods sections of the empirical studies.

4.2.2 Questionnaires

Rating Scales for User Experience

The swiftest means to collect a coarse measure for overall satisfaction is to present a simple *visual analogue scale* (VAS, 0-10) immediately after a BCI session, with 0 referring to "not satisfied at all". Holz et al. (2015a) introduced and established several additional items for use after every session: They included two additional VAS (*enjoyment* and *frustration*) and the three scales *level of exhaustion* (low, medium, high), *subjective level of control* (4 levels, low to high), and *control change* during the session (worse, same, better control).

NASA-TLX

The NASA Task Load Index was originally developed by NASA's *human performance group* to assess the subjective **workload** of a task and, further, to delineate the sources of said workload (Hart & Staveland, 1988). It has since been used in many areas, including BCI research (Baykara et al., 2016; Lorenz et al., 2014; Zickler et al., 2011). The TLX comprises six subscales that are split into mental, physical, and temporal demands to describe the source of the user's subjective effort, whereas performance, effort and frustration levels are used to describe the user's experience while working on the task. All items are rated on a 20-step unipolar rating scale (0-100, from low to high). By averaging their score, the *Raw TLX* can be calculated, and some researchers argue that this score is sufficiently meaningful on its own (Byers, 1989). In fact, the

Raw TLX is sometimes used in BCI studies as well (Lorenz et al., 2014). Nonetheless, the original authors included a second part into the NASA-TLX that includes pairwise comparisons of the six subscales to generate relative weights. Here, the participant chooses the subscale they consider more important to their workload. The scales of the Raw TLX are then multiplied with their respective weights to generate the overall NASA-TLX score as intended by Hart and Staveland.

The TUEBS (Extended QUEST 2.0)

The *Quebec User Evaluation of Satisfaction with assistive technology* (QUEST 2.0, Demers et al., 2002) was designed as a standardized tool to quantify **satisfaction** with an AT product.

All main items in the questionnaire consist of ratings on a Likert-type scale, from 1 ("not satisfied at all") to 5 ("very satisfied"). The participants are invited to comment when a rating was below 5. The total satisfaction score of the QUEST is given by the average of all rated main items. Additionally, the participants are asked to indicate the three most important factors for BCI use, out of twelve possibilities overall.

The studies included in the present work used an extended version of the QUEST 2.0, which was adapted for evaluation of BCI use and dubbed the TUEBS. For this adaptation, Zickler et al. (2011) removed four items that were considered inadequate or not useful in the context of BCI. The remaining items covered aspects such as the dimensions of the device, weight, safety, comfort, effectiveness, quality of service, as well as the ease of use and setup.

In exchange, four items specifically for BCI use were added: reliability, speed, learnability, and aesthetic design (for details, see Kübler et al., 2014; Zickler et al., 2011). The QUEST is still valid when at least 6 of the 12 original items are intact - however, added items should not contribute to the overall score (Demers et al., 2002). The TUEBS requires considerable time to complete, especially when used with patients, and should therefore be used only infrequently.

Still, all of the above questionnaires can be presented in certain intervals during long-term BCI use to explore possible effects of training on these subjective measures - after all, the UCD calls for the involvement of the end users in an iterative design workflow.

4.3 Efficiency and Training Effects of P300 BCI

4.3.1 BCI Usability

Although many users perform very well in their use of a visual BCI paradigm, even on their first try, an estimated 20 % of people can not achieve meaningful control over EEG-based BCIs (Allison & Neuper, 2010) - a phenomenon often described as *BCI-inefficiency* (Kübler et al., 2011).

In auditory or tactile BCIs, which are often more challenging to use (as compared to visual paradigms), this issue is magnified - both for healthy subjects and patients. This is usually reflected in substantially decreased BCI performances (Furdea et al., 2009; Käthner et al., 2013; Severens et al., 2014), likely because of an increased workload during the use of non-visual modalities (Halder et al., 2019). As a consequence, the prevalent criterion for successful BCI use of at least 70 % accuracy (Kübler et al., 2001b) is frequently not met.

4.3.2 Training Effects in P300 BCI

Current Examples

For SMR-based BCIs, training is well described and often considered necessary to utilize such paradigms to their full potential (Grosse-Wentrup & Schölkopf, 2013). A different situation presents itself for P300 BCIs: While there is evidence that training can generally affect ERPs (e.g. Kraus et al., 1995; Tremblay et al., 2001), studies that implement a training program with paradigms based on the P300 are scarce.

Still, training effects on P300 BCIs have recently come into the focus of some researchers, who aspired to improve overall performance of these paradigms. There are a few studies that could demonstrate a positive impact of repeated BCI use:

For instance, Baykara et al. (2016) demonstrated that, across five training sessions using an auditory P300 BCI, classification accuracies and ITRs had increased significantly ($N = 16$, healthy). Using the same auditory paradigm and five sessions, this effect was confirmed even for patients with motor impairments ($N = 5$), showing significant increases in P300 amplitudes and, again, in accuracies and ITRs (Halder et al., 2016).

Similar studies have also been performed for the tactile modality: In the study by Herweg et al. (2016) (the direct predecessor of the present work, further described in section 5.2.3), healthy, elderly participants attended five training sessions during which they used a four-class tactile P300 BCI for wheelchair navigation. Again, significant positive effects of training were discovered on offline ITRs and P300 amplitudes.

Such performance increases however, are not a matter of course, especially among impaired users. A study by Nijboer et al. (2008b) encouragingly showed that ALS patients were able to use the visual spelling matrix with stable performances over many months - but performance *increases* over time could not be demonstrated.

Possible Factors for Performance Increase

This begs the question what specifically is trained during BCI use, and which factors are responsible for performance increase. To date, these aspects are not fully understood.

One possible explanation is that training-induced improvements, especially in tactile paradigms, are related to **sensory perception**. Participants might become more sensitive toward the stimuli, leading to an improved target/non-target discrimination ability, which is important for BCI performance (Comerchero & Polich, 1999).

While a review of current literature provides no evidence for training-induced changes on the cellular level of the mechanoreceptors themselves, several studies have shown that somatosensory perception pathways can be trained:

For instance, Imai et al. (2003) observed a continuous performance increase during 30 days of intensive tactile frequency discrimination training ($N = 16$, healthy). The increase was strongest during the first two weeks of training, after that, only minor or no changes occurred. However, they found no increase in the cortical representation of the finger that was used for the task. A similar study on adult owl monkeys on the other hand provided evidence that training tactile discrimination is mirrored in the expansion of cortical representation (Recanzone et al., 1992). Further human studies seemed to confirm this, as researchers have found that, when blind people learn Braille, sensorimotor cortical areas representing the respective finger enlarge with training (Hamilton & Pascual-Leone, 1998).

Tactile learning can even affect higher-order processing and the resulting P300 ERPs (Reuter et al., 2014). Although participants only attended two training sessions, the authors reported improved tactile discrimination accuracies, reduced latencies and increased amplitudes of the P300. Notably, Reuter et al. (2014) also showed that participants of advanced ages (57-67 years) were not negatively affected in their training success, and that the relative success of tactile learning was mostly dependent on the baseline performance before training. Training-induced somatosensory changes however could not be conclusively demonstrated.

Secondly, some studies suggest another possible factor for performance improvement, which is the effect of training on mental **workload**. The fact that high workload has a negative effect on BCI performance is already well established. A number of BCI studies manipulated mental workload by introducing an additional, simultaneous task. For instance, an auditory P300 BCI study by Käthner et al. (2014) found that when participants had to perform a dichotic listening task simultaneously to the BCI, P300 amplitudes and BCI performances were significantly reduced. Similarly, in a study on a tactile BCI, Thurlings et al. (2013) had participants perform a concurrent visual monitoring task while operating the BCI and reported comparable results.

The theory that practice leads to a workload decrease due to *automatization* is supported by several functional imaging studies showing changes in the activation patterns of cortical networks after many repetitions of a certain task (for a review, see Kelly & Garavan, 2005). Studies commonly describe that performance of a novel task initially elicits a widespread cortical activation, which subsequently decreases with repeated practice. For instance, this has been shown for bimanual coordination tasks (Debaere et al., 2004), mirror reading (Poldrack & Gabrieli, 2001) or visual search tasks (Kübler et al., 2006). Often, the shift from widespread activation toward the increase of activation of task-related areas implicated in attention and retrieval is interpreted as an improvement of neural processing efficiency (Kelly & Garavan, 2005).

While none of the examples above included a BCI, some of the tasks and required skills are comparable to those of many BCI paradigms. Any insights may therefore be transferable, which is why these effects should be considered in multi-session BCI studies. Indeed, Halder et al. (2019) recently measured hemodynamic activity during the use of an auditory EEG-BCI paradigm to further investigate the underlying neural mechanisms of training effects. Here, ten healthy participants attended five BCI training sessions, while brain activity was recorded with fMRI during the first and last session. In the hence possible pre/post comparison, the authors discovered that training had led to a decreased activity in the *superior frontal gyrus* and an increased activity in both the *superior temporal gyrus* and *supramarginal gyrus*, a finding that they attributed to improved stimulus perception and a reduction of mental workload.

The observations in this study were interpreted in the context of Logan's *instance theory of automatization* as a possible mechanism for workload reduction (Logan, 1985, 1988). According to this model, automatization is the gradual transfer of a mental algorithm for handling a novel task into an acquired, memory-based solution, driven by repeated practice. Performing a task under consistent conditions would therefore generate a memory representation, or *instance*, which strengthens with every repetition. Thus, while a certain task is initially handled with a potentially workload-intensive algorithm, later task performance would become gradually more *automatized* and handled by memory instead. This automatized solution is assumed to be less resource-intensive and might completely substitute the original algorithm after a certain period of training.

5

Implementations of P300 BCI

Having established this theoretical background, we can now have a closer look at two important examples of P300 based BCI paradigms: The visual spelling matrix and a tactile BCI for wheelchair control.

5.1 The Visual P300 Spelling Matrix

Although the visual modality was not used in this thesis, the concept of P300-based BCI paradigms is best explained with the example of the *P300 spelling matrix*. The matrix was established by Farwell & Donchin (1988), is very easy to use and prevalent up to this day.

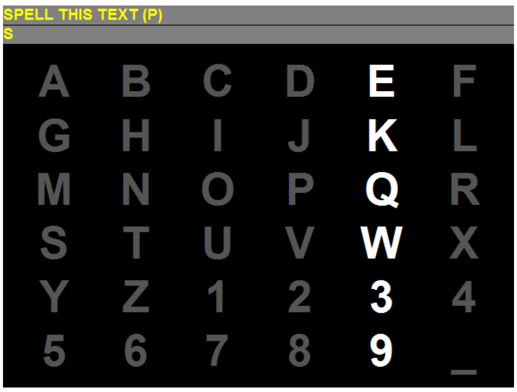

Figure 5.1: The P300 spelling matrix as implemented in BCI2000. During stimulation, entire rows and columns are highlighted (here, the column starting with E). If either contains the target letter, a P300 is elicited. Feedback can be displayed above the matrix.

In this derivation of the visual oddball paradigm, the user is presented with a 6x6 matrix of selectable characters (see Figure 5.1). Stimulation is performed by the random *flashing* (i.e. highlighting) of the characters - however, individually flashing all 36 characters would take rather long and would shift the target/non-target-ratio too far away from the desired 0.2 (Cohen & Polich, 1997). Instead, entire rows and columns are flashed, while the user still concentrates only on the desired character. Thus, if a row or column contains their target, a P300 will be elicited, but the precise identity of the character needs to be inferred: Using the temporal concurrence of stimulation and ERP elicitation, the BCI simply identifies the row and column based on P300 detection and then selects the character at their intersection. With this approach, only twelve events are necessary to cover the entire matrix and all possible selections of this 36-*class*-paradigm. The process of running through all possible events exactly once is commonly referred to as one *sequence*. As established in section 3.4, usually more than one sequence is necessary to collect enough data to support an accurate classification. Finally, the selected character is displayed as feedback on top of the matrix and the next stimulation phase begins. This entire process can be reviewed in the manual for BCI2000 (Schalk & Mellinger, 2010).

The visual matrix often allows for fast and reliable spelling with nearly 100 % accuracy among motivated, healthy users (Kleih et al., 2010). Impaired users may experience a diminished BCI-efficiency. For those, the visual spelling matrix has recently been adapted to overlay the characters with small pictures of a human face instead of simply highlighting them. Face recognition elicits the N170 ERP, which can contribute to classification (Kaufmann et al., 2011). Further, the spelling matrix can be used for a variety of applications other than mere communication, for instance as a creative tool in the case of *Brain Painting*[1], in which simple drawing commands are selected instead of letters (Botrel et al., 2017; Münßinger et al., 2010).

Although this visual paradigm is both reliable and well established, there are some limitations that should be discussed.

[1] www.brainpainting.net

5.2 A Case for Non-Visual Modalities

5.2.1 Dependency on Gaze and Vision

A recurring issue when using BCI with ALS patients is the fact that many paradigms, for instance the P300 spelling matrix introduced above, rely largely on the visual channel for both stimulation and feedback. Thus, they require that vision and gaze control be intact to at least some degree. As outlined in chapter 2, certain patients, especially in late stages of ALS, may experience these limitations and are likely to face usability issues with visual systems. Indeed, it has been shown that such paradigms' performances can suffer substantially when gaze control is restricted (Birbaumer & Cohen, 2007; Brunner et al., 2010).

Another argument for gaze-independent designs is the desire to leave the visual channel available for social interactions and other daily tasks, with the intent to not further restrict the patients' habits. Tactile stimulators may even be hidden beneath the user's clothing and completely escape the notice of others (Brouwer & Van Erp, 2010).

Because of these reasons, and with the potential end-users in mind, many alternatives to vision dependent BCIs have come into the focus of current research (for a review, see Riccio et al., 2012).

5.2.2 Gaze-Independent Alternatives

There is a number of approaches that have tried to solve the issue of gaze-dependency while still using the visual modality.

For instance, Treder & Blankertz (2010) developed the *Hex-o-Spell* as an alternative to the visual spelling matrix. This paradigm does not require a precise focusing of the target letter, but relies on covert attention on much larger target areas that consist of six flashing discs arranged around the centre of the screen. Once a disk has been selected, each of the six letters contained within is presented in a second, identical step.

Another approach eliminates gaze-reliance entirely, and simply applies colour stimulation through closed eyelids (Hwang et al., 2015). In this study, healthy participants were able to answer simple questions with a mean accuracy of 74.6 % by concentrating on one of three different colour stimuli. Stimulation was applied via LEDs which were attached to glasses worn by the participants.

Moreover, several non-visual BCI modalities have recently been explored. One approach that requires no intact gaze control is based on SMR (introduced in section 3.3), which can be modulated by the user via motor imagery. SMR-based BCI has already been successfully used for virtual wheelchair control (Velasco-Álvarez et al., 2011). Although SMR-BCIs were demonstrated to be generally feasible, some studies concluded that they sometimes still lack efficiency or are too unreliable in their performance (Guger et al., 2003; Hammer et al., 2012).

On the other hand, many ERP based BCI paradigms have been developed which use auditory (Baykara et al., 2016; Halder et al., 2016; Hill & Schölkopf, 2012) or tactile stimulation (Herweg et al., 2016; Kaufmann et al., 2014; Waal et al., 2012). For instance, in the study by Halder et al. (2016), the authors used their auditory BCI which encoded rows and columns of the spelling matrix with animal sounds. They achieved mean accuracies of 52.8 % and ITRs of 3.08 bits/min among motor-impaired patients. Similarly, Baykara et al. (2016) reported an average ITR of 5.14 bits/min, using the same paradigm with healthy participants. Hill & Schölkopf (2012) developed a novel approach, for which they split the stimuli in two simultaneous audio streams applied to the left and right ear, each stream containing either targets and non-targets of the two possible selections "yes" and "no". With this paradigm, they achieved a mean accuracy of 85 % with an ITR of 4.98 bits/min. These studies, along with a few further examples of auditory BCIs, are summarized in Table 5.1.

As for the tactile modality, some studies successfully used Braille stimulators or similar devices that applied small taps at the fingertips of the participants: Waal et al. (2012) repurposed the 6x6 matrix that is usually used in visual paradigms, but applied Braille stimulation instead of flashes. Using this paradigm, they reported a mean 6-class classification accuracy of 67 %. A study by Severens et al. (2014) used this paradigm and achieved 56 % and 53 % in healthy and ALS patients, respectively. Vibrotactile stimulation has also been used for BCI, for instance in a study by Brouwer & Van Erp (2010), where up to six devices on a belt around the waist of healthy participants were used to apply short vibrations. Here, accuracies of up to 73 % were achieved. Vibrotactile stimulation was also shown to work in ALS patients in a study by Silvoni et al. (2016), who implemented a simple tactile oddball task. Interestingly,

Table 5.1: Overview of auditory BCI studies mentioned in the text.

Publication	Accuracy Mean [%]	ITR Mean [bits/min]	Classes	Population	Analysis
Kübler et al. (2009)	12.10	0.05	25	ALS	On
Höhne et al. (2011)	89.4	4.61	9	H	On
Schreuder et al. (2011)	77.0	5.27	6	H	On
Hill & Schölkopf (2012)	85	4.98	2	H	On
Halder et al. (2016)	52.8	3.08	25	MI	On
Baykara et al. (2016)	77.9	5.14	25	H	On
Onishi et al. (2017)	84.7	-	2	H	Off
	90.0	-	2	ALS	On
Kaongoen & Jo (2018)	95.61	2.97	2	H	Off

Notes: Classes = Number of possible outcomes, H = Healthy, MI = Diverse motor impairments, On = Online, Off = Offline

no significant evidence for ERP pathology was found in a sample of 14 ALS patients. Despite slightly smaller P300 amplitudes and decreased classification accuracies (as compared to a healthy control group), no correlation with the clinical ALS stage was revealed, and the authors concluded that a BCI application among this user group was feasible.

One tactile paradigm by Kaufmann et al. (2014) and Herweg et al. (2016) has especially inspired the present work. Due to very encouraging results from these studies, a more thorough look at the paradigm is merited.

5.2.3 A Tactile BCI for Wheelchair Control

All studies included in this thesis are based on the following four-class vibrotactile paradigm, which was designed and proven to be feasible for wheelchair control by Kaufmann et al. (2014) and Herweg et al. (2016). In the 2016 study, nine healthy, elderly participants (ages ≥ 50) used the system to navigate a virtual wheelchair through a small 3D apartment across five training sessions. The chair could be controlled using the BCI by selectively concentrating on the stimulus that encoded the desired movement direction (details on the paradigm are elaborated in the general methods, chapter 7). Feedback was given in form of the wheelchair executing the classified command on a monitor (Figure 5.2). This visualization of wheelchair navigation, although simple, proved to be adequate to enable all participants to achieve high levels of BCI performance: During session five, participants achieved a mean accuracy of 92.56% and mean ITR of 4.98 bits/min.

Figure 5.2: The virtual environment from the user's perspective. Screenshot was taken from a slightly modified version from Herweg et al. (2016).

Most notably however, Herweg et al. (2016) introduced an optional task at the end of session five. Here, the experimenters reduced the number of stimuli based on the participants' individual capabilities, who then achieved a mean ITR of 20.73 bits/min (mean accuracy 95.56 %) - a value that is still the highest ever reported from a tactile BCI - for comparison, Table 5.2 provides an overview of recent tactile BCI studies based on P300 paradigms. These surprisingly high ITR and accuracy values aside, the study also demonstrated that the tactile paradigm was trainable (section 4.3.2), even for participants in an age range similar to potential end-users. Significant training effects could be shown for P300 amplitudes (5.46 to 9.22 μV), but not for online accuracies or ITRs, a fact the authors attributed to a ceiling effect due to accuracy values already close to 100 % during the first session.

A possible factor for the paradigm's success may have been the fact that the stimuli encoding the four possible wheelchair commands were applied to body positions which were easily associated with the desired direction for movement. For instance, left and right knees encoded left and right turns, respectively, therefore requiring little or no memorization of their meaning. About two years before, Thurlings et al. (2012b) had already demonstrated that congruence of directions had a positive effect on performance. Other preceding studies often applied the stimuli to the fingertips, but reported that participants found it difficult to discriminate between the different positions, which were relatively close the each other (Severens et al., 2014; Waal et al., 2012). In the paradigm by Kaufmann et al. (2014) and Herweg et al. (2016) on the other hand, the vibrotactile devices were spaced far apart on opposite sides of the body, likely facilitating the discrimination between stimuli (Brouwer & Van Erp, 2010).

Table 5.2: Overview of recent tactile P300 BCI studies. If variations of paradigms were reported, those with the highest ITRs are included here. If numerical values were not explicitly reported, approximations based on plots are provided. Hybrid BCIs that used a combination of modalities are not included.

Publication	Accuracy Mean [%]	ITR Mean [bits/min]	Classes	Population	Analysis
Brouwer & Van Erp (2010)	68[a]	3.71	6	H	On
Waal et al. (2012)	67.0	7.8	36	H	Off
Thurlings et al. (2012a)	78[a]	6.52	6	H	Off
Ortner et al. (2012)	68.1	3.8	8	H	On
Severens et al. (2013)	77.0	1.2	2	H	On
Severens et al. (2014)	60.0	6.6	36	H	Off
	58.0	6.6	36	ALS	Off
Kaufmann et al. (2014)	85.8	2.54	4	H	On
Herweg et al. (2016)	**95.6**	**20.73**	**4**	**H**	**On**
Halder et al. (2018)	71.0	3.4	5	H	Off

Notes: [a] = Approximation, Classes = Number of possible outcomes, H = Healthy, On = Online, Off = Offline

6

Research Goals

Since the tactile BCI by Kaufmann et al. (2014) and Herweg et al. (2016) (section 5.2.3) had already been demonstrated to be not only feasible, but also to enable fast and reliable operation, we decided to continue research and development of this paradigm, as guided by the following major goals:

- **Goal 1:** The previous results, although very encouraging, were limited by a very small sample size and a putative ceiling effect. Therefore, we aimed to replicate this study to provide further support for the paradigm's training effects and excellent performance (Study I and Study II).
- **Goal 2:** Assuming that training effects of the tactile paradigm could be confirmed, we wanted to learn more about their physiological origins (Study II). This knowledge could later prove valuable to optimize the paradigm.
- **Goal 3:** Our research is motivated by the desire to assist the actual end-users. Therefore, it was crucial to involve this group in the research progress, so we intended to demonstrate the paradigm's feasibility among patients with ALS. Further, we wanted to learn about effects of training and disease progression in ALS users (Study III).

Part II of the thesis will now provide a detailed description of three empirical studies which were designed and conducted in order to approach those goals.

Part II

Empirical Investigation

7

General Approach

The tactile BCI paradigm was fully adopted from Herweg et al. (2016). For ease of comparability and internal consistency, care was taken to keep the design of the paradigm and the tools for assessment and analysis identical between the following studies. Hence, to avoid redundancies, certain shared features such as electrode positions and details of the tactile stimulation are reported here. Unless indicated otherwise, all studies in the thesis are based on this design.

7.1 Tactile Stimulation

Tactile stimulation was applied as short vibrations at the participant's right and left thigh (3-5 cm above the knee), abdomen (2-5 cm above the navel) and upper back (at the height of the C4–T3 dermatomes) via BCI2000-controlled devices called *tactors*. A schematic of the tactor placements is shown in Figure 7.1. Each position encoded a congruent direction, e.g. the stimuli on the abdomen were associated with the *forward* command, which slowly moved the wheelchair forward approximately 1.5 meters. The position on the user's back encoded the *backward* command. Stimuli on the left and right knees encoded the command to turn 45 degrees left and right, respectively.

Tactors were attached using Velcro bands or adhesive tape and adjusted until vibrations at all positions were reported as equally strong by the participants. During the stimulation phase of the BCI, all four tactor positions were activated individually (vibrating for 220 ms) in a pseudorandomized order with equal probabilities of 25%. The interstimulus interval was 400 ms.

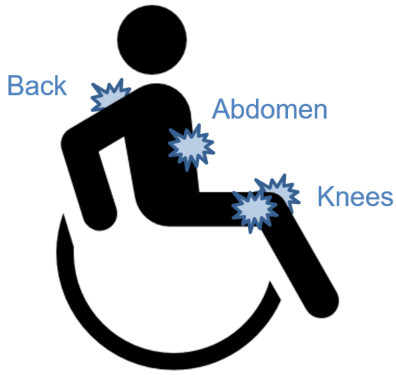

Figure 7.1: Schematic of tactor positions on a potential user. Positions at the abdomen, back, left and right knees encoded the movement commands *forward, backward, left* and *right turn*, respectively.

For participants in the Würzburg lab, the "C2 tactors" by *Engineering Acoustic Inc.* (Casselberry, USA) were used (see Figure 7.2). The tactors were set to 250 Hz and maximum intensity. Patients at Aalborg University or at their home used a custom-made prototype based on an Arduino microcontroller (*Adafruit HUZZAH32 ESP32 Feather*, shown in Figure 7.2). This tactor device was developed in collaboration with the Technische Hochschule Nürnberg Georg Simon Ohm (Schuster, 2018) and designed to be fully interchangeable with the C2 tactor system. Hence, both devices could be controlled identically by an unmodified BCI2000, requiring only C++ or Python 2.7 based adapter programs. Since BCI2000 does not natively support the control of third-party hardware and software, communication to the tactile devices, as well as to the virtual environments, was established via the *User Datagram Protocol* (UDP), over which BCI2000 broadcasted all events of the BCI cycle.

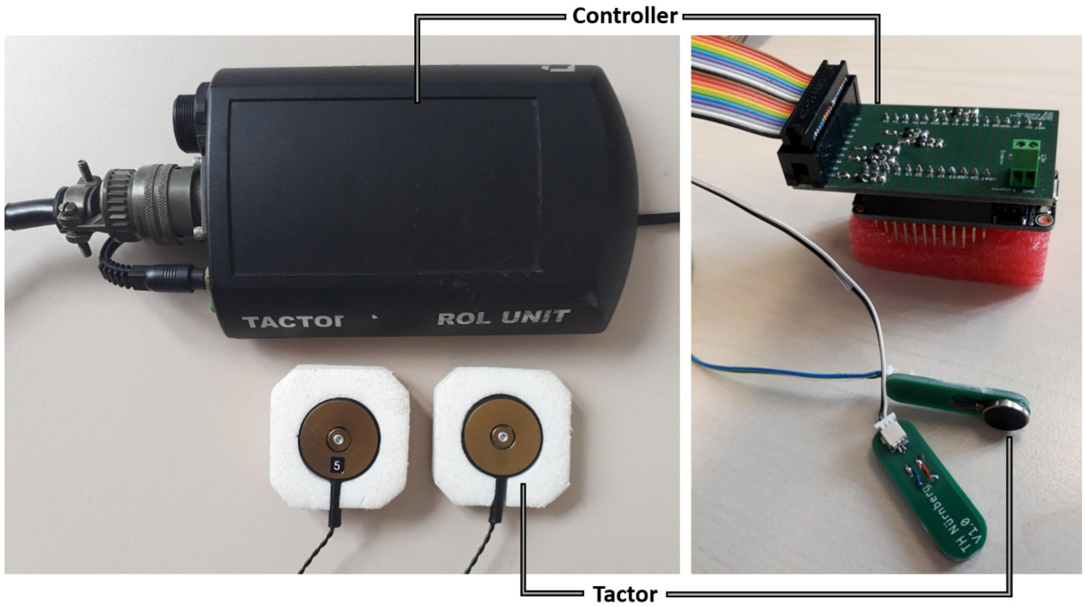

Figure 7.2: Tactor devices. **Left:** C2 tactors by Engineering Acoustic Inc. **Right:** Prototype by the Technische Hochschule Nürnberg, based on an Arduino microcontroller.

7.2 EEG Recording and Processing

EEG data at the University of Würzburg was recorded from 12 passive Ag/AgCl electrodes (*g.tec Engineering GmbH*, Graz, Austria). Patients at their homes or at Aalborg University used an active electrode system (*g.GAMMASYS*). The signal was amplified using a *g.tec g.USBamp* and recorded with a sampling rate of 512 Hz in either case. Impedances were kept below 5 kΩ. Electrode positions according to the 10-20 system (Jasper, 1958; Sharbrough, 1991) were Fz, FC1, FC2, C3, Cz, C4, P3, Pz, P4, O1, Oz, and O2. For recordings with passive electrodes, ground and reference electrodes were placed at the right and left mastoids, respectively. In the case of the active (*g.GAMMASYS*) electrodes, AFz and the left earlobe were used as ground and reference.

The BCI cycle was controlled with BCI2000 (Schalk et al., 2004). Classifier weights were defined using the step-wise linear discriminant analysis as implemented in the BCI2000 package. EEG data was filtered online with a notch filter between 48 and 52 Hz and a band-pass filter between 0.1 and 60 Hz.

Offline data processing was performed with MATLAB©(v2013b) using adapted scripts provided by the BCI2000 package and EEGLab (Delorme & Makeig, 2004).

For plot generation and statistical analysis, EEG data was band-pass filtered between 0.1 and 30 Hz. Segments from -100 ms to 800 ms around the stimulus onsets were created. The pre-stimulus epoch from -100 ms to 0 ms was used for baseline-correction. Segments containing extreme values exceeding a voltage threshold (e.g. $\pm 100\ \mu V$) were rejected. Target and non-target segments were averaged separately for feature extraction to create single-subject averaged EEG responses. Grand averages were obtained by averaging the segments over all participants.

To evaluate the P300 from single-subject averages, we extracted the *mean P300 amplitudes* from a relevant P300 time window (300-500 ms) of the target segments. This was achieved by averaging all samples in the window into one mean value per participant.

As a measure of target/non-target divergence, and to account for possible changes of the non-target curve, we also extracted the *mean difference between target and non-target*. This was done by generating the difference curve by subtracting non-target curve from the target curve. This difference curve was then averaged in the same fashion as the mean P300 amplitude. Both measures entered analysis as dependent variables. For better readability, the term *mean difference between target and non-target* will be abbreviated as *mean difference* or *difference between curves* from here on.

7.3 General Session Procedure

Participants were sitting in a chair in front of a monitor in a distance of approximately 1.2 m. Patients who could not sit in a regular chair were sitting in their own specialized wheelchair or a similar mobility aid. Participants were asked to keep their eyes open and to avoid unnecessary movements or blinking and to keep all facial muscles relaxed during BCI recordings.

After EEG preparation and tactor placement, three BCI calibration runs were performed (see section 3.5.3). The calibration procedure was repeated every session to account for possible changes due to training. Participants were instructed to concentrate on specific target stimuli. To help focus attention on the targets, mentally counting their occurrence was suggested as a possible strategy. All other (non-target) stimuli had to be ignored. The fixed order of targets/commands for one calibration run was *left, right, forward, backward, left, right, forward, backward.* One run thus comprised eight direction commands, with every direction/body position being the target twice. The target was changed to the next position after ten sequences. The current target was always indicated on the monitor.

Overall, three calibration runs generated a total of 240 target and 720 non-target epochs available as input for machine learning. These data were used to train a linear classifier using swLDA.

Next, the resulting classifier was used to provide online feedback during one or more BCI tasks, depending on the study. Those were either variants of the calibration task, with the addition of online classification and feedback, or virtual wheelchair navigation. Since all tasks were different, their specifics are reported in the methods section of the respective studies. Participants could take a short break between all tasks.

7.4 Data Analysis

7.4.1 General Statistical Approach

Data derived from online BCI performance measures, offline processed EEG, and questionnaires were subjected to statistical tests. The approach was largely guided by the recommendations of Bortz & Schuster (2011). In cases of small sample sizes or violations of the assumptions of parametric tests (e.g. normal distribution), non-parametric tests were used and reported.

The threshold for statistical significance (*) was defined at an alpha error level of 5 % ($p < .05$). Further, p-values $< .01$ were considered as highly significant (**). Due to small sample sizes, we refrained from adjusting alpha levels for multiple comparisons (between sessions), as suggested by Nakagawa (2004). Therefore, Bonferroni-corrections were performed only if explicitly stated. Instead, we focused on hypothesis-driven, planned contrasts, for instance pre/post comparisons of only the first and last sessions of training, while reporting effect sizes using the *partial eta squared* ($\eta\mathrm{p}^2$, for ANOVAs) or Cohen's d_z (Lakens, 2013) for t-tests. Distributions around observed mean values are reported using the *standard deviation* (SD). In plots, the *standard error of the mean* (SEM) is indicated instead, as the SD was often too large to display.

Statistical calculations were performed with SPSS statistics 25 (IBM, USA).

7.4.2 Cross-validation

For offline accuracy calculations, a simple leave-one-out cross-validation approach was used to prevent overfitting of the model (Stone, 1974): Out of N available EEG datasets, *N-1* were used to train the classifier weights, which were then tested on the one remaining data set to calculate the accuracy. After this was done N times to cover all possible permutations, the average accuracy value entered the statistical analysis.

7.5 Ethical Approval

Studies I and **II** were considered replications or direct continuations of the study by Herweg et al. (2016), which was approved by the ethical review board of the Institute of Psychology at the University of Würzburg, Germany (GZEK 2013-11). Beyond that, all studies were performed in accordance with the declaration of Helsinki (General Assembly of the World Medical Association, 2014).

8

Study I: Training Effects of Wheelchair Control in a Virtual Environment

8.1 Introduction and Motivation

Our first goal was to continue and extend the work by Herweg et al. (2016), which has already been outlined in the general introduction (section 5.2.3). Briefly, its most important achievements were:

- Novelty of a tactile P300 paradigm for wheelchair control
- Exceptionally high BCI performance
- Significant evidence for training effects across five sessions
- A healthy, but age-matched sample (to potential end-users)

Because of these merits, we decided to perform a thorough replication of this study to further solidify the evidence of the feasibility, high performance and trainability of the tactile BCI paradigm.

8. STUDY I: TRAINING EFFECTS OF WHEELCHAIR CONTROL IN A VIRTUAL ENVIRONMENT

Notably, the previous study had a number of limitations which we wished to address:

- A very small sample size ($N = 9$)
- A putative ceiling effect
- The lack of any subjective data (e.g. usability ratings)

Thus, our goal for this replication study was not only to recruit a larger sample, but also to find a way to avoid ceiling effects that may have occurred in the previous study. Further, we wanted to include questionnaires to collect subjective data related to the UCD concept. In all other regards, the tactile paradigm and overall study design was kept identical.

The following hypotheses were devised:

H1: Based on previous results, we hypothesized that physiological EEG measures, specifically *P300 amplitude* **(H1a)** and *difference between curves* **(H1b)**, would significantly increase across five training sessions.
H2: Similarly, we expected that ITRs during online BCI would increase significantly with training.
H3: Furthermore, we hypothesized that five training sessions would be sufficient for all participants to achieve at least 70 % accuracy during wheelchair navigation in their last session.
H4: Because of the excellent performance of elderly participants in the study by Herweg et al. (2016), we expected that participants would not be negatively affected by their advanced age, as compared to young adult participants.
Exploratory: Additionally, we added several questionnaires about usability, i.e. workload and general satisfaction, to explore relevant factors for future improvements of the paradigm.

8.2 Methods

8.2.1 Participants

This study was conducted at the University of Würzburg, Institute of Psychology (Intervention Psychology). Sixteen healthy participants were recruited over social media and notice boards at the University. We made an effort to invite several middle-aged participants to allow for a comparison with the study by Herweg et al. (2016) (which was performed with participants over 50 years of age), and to test for potential age-related effects (H4). One participant did not wish to continue participation after the second session because of a conflict with their personal schedule. Therefore, $N = 15$ participants were included in the study (3 male, age range 20-61 years, $M = 38.0$ years, SD = 15.4). Of those, $N = 8$ comprised a group of young adults (ages 20-33 years, $M = 24.8$, SD = 4.4, 1 male), whereas $N = 7$ belonged to a middle-aged demographic group (ages 41-61 years, $M = 53.1$, SD = 8.1).

All participants were naïve with respect to BCI operation and reported normal or corrected-to-normal vision. A monetary compensation of EUR 7.50 per hour was paid. Alternatively, local students of psychology could receive course credits (participation in experiments is a part of the psychology degree). All read and signed informed consent forms about the procedure. The procedure was also explained verbally at the beginning of the first session.

8.2.2 Procedure

To investigate training effects of BCI use, participants were invited to five sessions on separate days, with one to seven days between subsequent appointments. Tactile stimulation was implemented using the C2 tactors as described in section 7.1. EEG was recorded and processed as outlined in section 7.2. After the three calibration runs at the beginning of each session, the resulting swLDA feature weights served to classify the commands during an online wheelchair navigation task in a virtual apartment. At the end of each session, participants were asked to fill in a 0-10 satisfaction scale and a digital version of the NASA-TLX. The TUEBS was applied only in the first and last sessions, since it took considerable time to fill in, and short-term effects were not expected for most of its items.

8.2.3 Virtual Environment

The virtual environment was implemented with Python 2.7 using a plug-in for 3D visualization (Panda3D[1]) by Kaufmann et al. (2014) and Herweg et al. (2016). For this replication, only minor changes regarding its visual appearance were made: The headrest of the wheelchair model was removed for improved visibility of the area directly in front of the wheelchair. Further, some simple graphic textures were introduced into the 3D model of the apartment.

Navigation was visualized on the monitor from a perspective similar to that of a person sitting in a wheelchair (see Figure 8.1). However, parts of the wheelchair were shown from behind to better allow for an estimation of its dimensions. The apartment consisted of a hallway and several interconnected rooms. To facilitate orientation, a small window on the top right on the monitor showed the apartment from a top-down perspective. The navigation program was connected to BCI2000 via UDP, enabling it to interpret and execute the BCI classification results as wheelchair movements - thus providing the feedback of this paradigm.

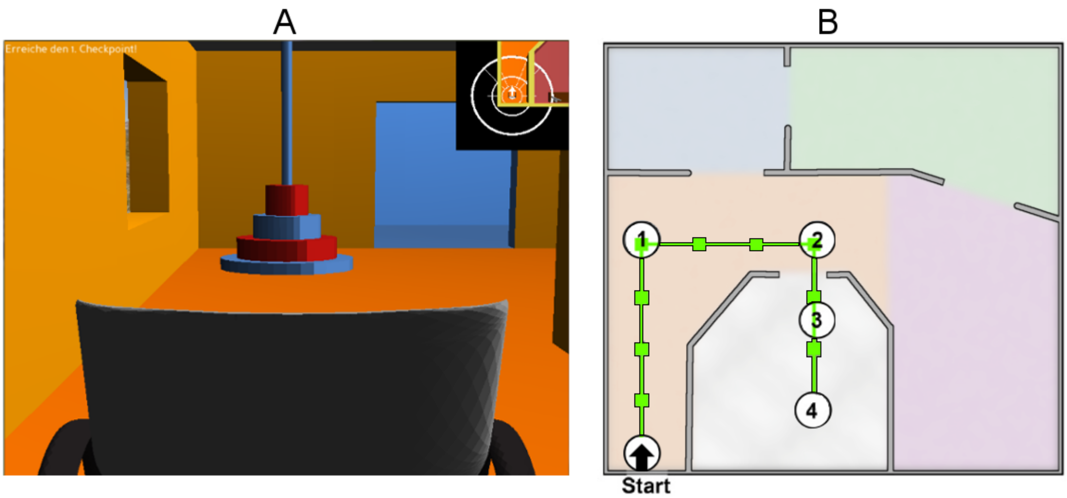

Figure 8.1: Navigation in the virtual apartment, adapted from Herweg et al. (2016); Kaufmann et al. (2014).
A: User's perspective, showing the backrest of the wheelchair and the first checkpoint in front. **B:** Top down schematic of the apartment illustrating one of the courses. Numbers indicate checkpoints, notches indicate travel distance for one movement command.

[1] www.panda3d.org

8.2.4 Navigation Task

In every session, two wheelchair navigation runs were performed. For this task, participants were asked to navigate along a certain course through the apartment which was outlined by four checkpoints that had to be reached sequentially in the given order (see Figure 8.1 B).

Participants were instructed how to control the wheelchair before they started. Successful control required them to first choose an appropriate direction command, and then concentrate on the stimuli on the corresponding body position during a tactile stimulation phase. For instance, if they intended the wheelchair to turn 45 degrees to the left, they were instructed to focus on the stimuli on the left knee. After the stimulation phase, the BCI classified the user's command based on the elicited ERPs.

Finishing the course required at least 14 correct commands, however, erroneous or misclassified selections were also accepted and executed, as this was a free navigation paradigm. Consequently, an erroneous command had to be either corrected or integrated into an alternative route toward the next checkpoint. Impossible commands, for instance when moving into an obstacle, were not executed but counted as an error. The maximum number of commands for one run was set to 22, which was estimated by Herweg et al. (2016) to be sufficient assuming an accuracy of 70 %. If participants had not reached the final checkpoint by then, the run was terminated. After each run, start and end points of the course were switched, so that participants had to navigate back to their original starting point. Additionally, the order of the courses was balanced between sessions and participants, so that not all participants started with the same configuration.

Importantly, unlike in the study by Herweg et al. (2016), who used a fixed number of eight sequences per command, we individually adjusted the number of sequences for every participant and session. This was done to pre-empt a ceiling effect of online performance as reported by Herweg et al. (2016). As a consequence, participants whose predicted classification accuracy was particularly high received a number of sequences smaller than eight, thus enabling them to select a command in less time due to the hence shortened stimulation phase. We configured the BCI to use the minimum necessary number to reach 100 % accuracy based on predictions from the classifier algorithm

(BCI2000's *P300Classifier.exe*) for every session prior to the navigation task. However, to avoid long selection times, the number of sequences was never set higher than ten, even when participants were not predicted to reach 100 %.

8.2.5 Statistical Analysis

Physiological measures were extracted from calibration data, as data from free BCI use (in this case, the navigation runs) do not contain the necessary identifiers of target/non-target segments. Preprocessing was performed as described in section 7.2 using an artifact threshold of ± 150 μV and a time window of 300-500 ms post-stimulus to extract mean amplitudes and mean differences. As in Herweg et al. (2016), we focused on positions Fz, Cz and Pz for analysis.

Offline accuracies were cross-validated over the three calibration runs using the maximum number of sequences (10) available for classification. The average online ITRs from the two navigation tasks of each session were calculated as described in section 3.5.2. Online accuracies from navigation are reported descriptively only, because accuracies are not comparable when the number of sequences between sessions or participants varies (for instance, a lack of a significant accuracy change might suggest there was no training effect on classification quality - when in fact, accuracies were artificially kept stable by reducing the number of sequences).

All other EEG or BCI measures were entered into separate Friedman ANOVAs. This non-parametric test was chosen as an alternative to the repeated measures ANOVA, which was considered not applicable because of small sample sizes and violations of the assumption of normal distribution.

To test for age-related effects on BCI efficiency, we split participants in two groups, young adult and middle-aged. Online ITRs from both groups were compared using a Mann-Whitney U-test.

8.3 Results

Fifteen participants completed all five BCI training sessions and thus provided a total of 75 BCI sessions available for analysis. Nine participants reached BCI efficiency (efficiency was assumed when mean navigation accuracies of 70 % or higher were reached during the last session, or in at least three other sessions). Participants 6, 7, 8, 10, 11 and 16 did not meet this criterion. As we wanted to adequately describe effects of BCI training, we excluded those participants from all relevant statistical tests and from grand average EEG plots. Since we considered subjective data from questionnaires informative irrespective of BCI efficiency, their analysis was performed on the full sample.

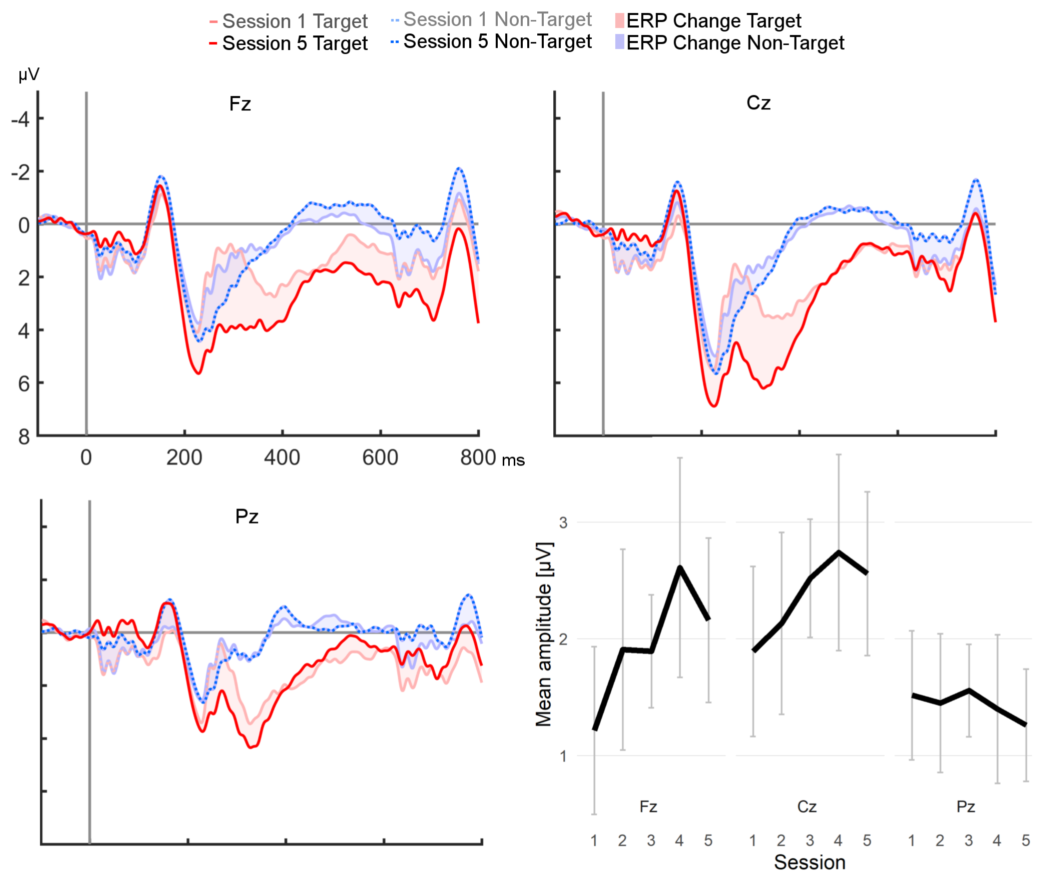

Figure 8.2: Grand averages showing target and non-target epochs from the first and last sessions at the midline electrodes (shaded areas between the curves visualize changes from session one to five). *Bottom right:* Extracted target amplitude values from all sessions (error bars represent SEM).

8.3.1 Physiological Measures

We plotted post-stimulus epochs of target and non-target curves at Fz, Cz and Pz from session one and five to visualize their development with training (**H1**). Figure 8.2 shows the grand average of all BCI efficient participants. Highest peak values (6.21 μV) were observed at position Cz in session five. Descriptively, the extracted ERP features (amplitudes and difference between curves) increased at Fz and Cz in the P300 range. However, Friedman ANOVAs did not reveal any significant effects on either measure over the course of the study. Planned comparisons (session 1 vs. 5; Mann-Whitney test) were not significant (all U-test $p > .1$). Table 8.1 provides an overview of descriptive statistics and Friedman ANOVAs.

Table 8.1: Statistics for ERP measures at Fz, Cz and Pz. No significant increases were supported by Friedman ANOVAs.

		Fz	Cz	Pz
Amplitudes	Mean S1 (SD) [μV]	1.92 (3.26)	2.73 (3.13)	2.14 (2.35)
	Mean S5 (SD) [μV]	3.10 (3.08)	3.94 (2.56)	2.36 (1.48)
	Chi-square (X^2)	1.96	5.33	.44
	Significance (p)	.769	.261	.985
	Effect size (d_z)	.65	.38	.18
Difference Between Curves	Mean S1 (SD) [μV]	1.48 (4.07)	2.32 (3.30)	2.14 (2.26)
	Mean S5 (SD) [μV]	2.70 (4.05)	3.50 (2.70)	2.38 (1.77)
	Chi-square (X^2)	5.96	2.04	1.51
	Significance (p)	.210	.754	.848
	Effect size (d_z)	.58	.65	.14

Averaged epochs of each participant are shown in Figure 8.3 for position Cz, at which mean amplitudes were highest. Visual analysis revealed no clear divergence of target/non-target curves among BCI inefficient participants, and a large divergence in participants 4, 13, 14 and 15. Notably, we observed a deflection of an inversed polarity in participant 9 (who achieved efficient control nonetheless). Fz and Pz are shown in the supplementary Figures A.1 and A.2.

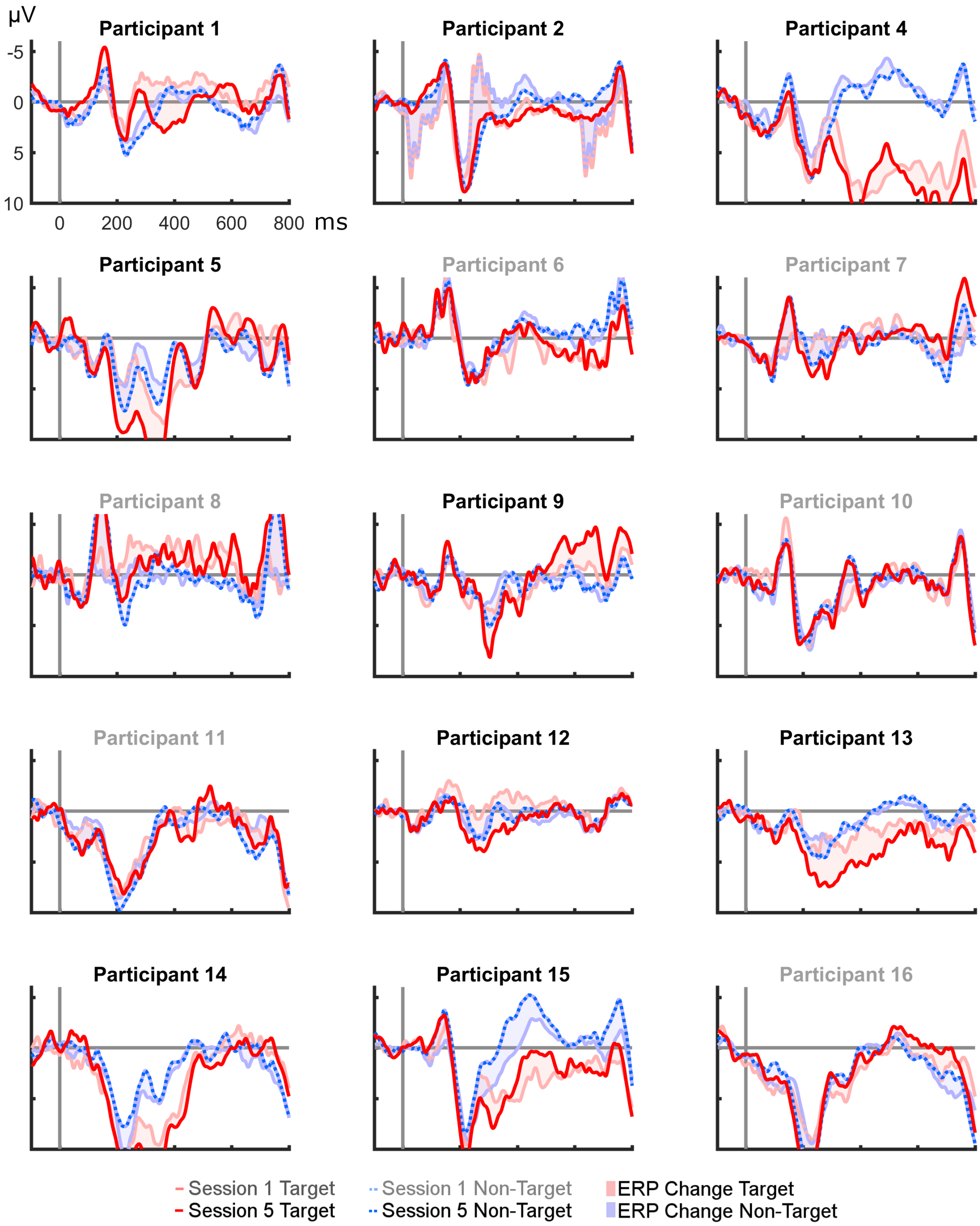

Figure 8.3: Individually averaged post-stimulus epochs at Cz of all participants' first and last sessions. Shaded areas indicate changes. Inefficiencies are indicated by a gray participant code.

8.3.2 BCI Performance

To test for training effects on BCI performance (**H2**), we calculated three separate Friedman ANOVAs. Their results are summarized in Table 8.2. Across the five training sessions, online accuracies increased significantly from 65 % to 86 % ($p < .05$) and online ITRs increased highly significantly from 3.10 to 9.50 bits/min ($p < .01$) (Figure 8.4). Similarly, cross validated offline accuracies increased highly significantly between sessions, from 70 % to 95 % ($p < .01$).

These values greatly exceeded 50 % which is the approximate chance level threshold to be considered non-random and statistically significant (using an alpha of .01; Combrisson & Jerbi (2015); Müller-Putz et al. (2008)). Figure 8.4 shows the development of ITRs and accuracies of the BCI efficient participants across the five training sessions.

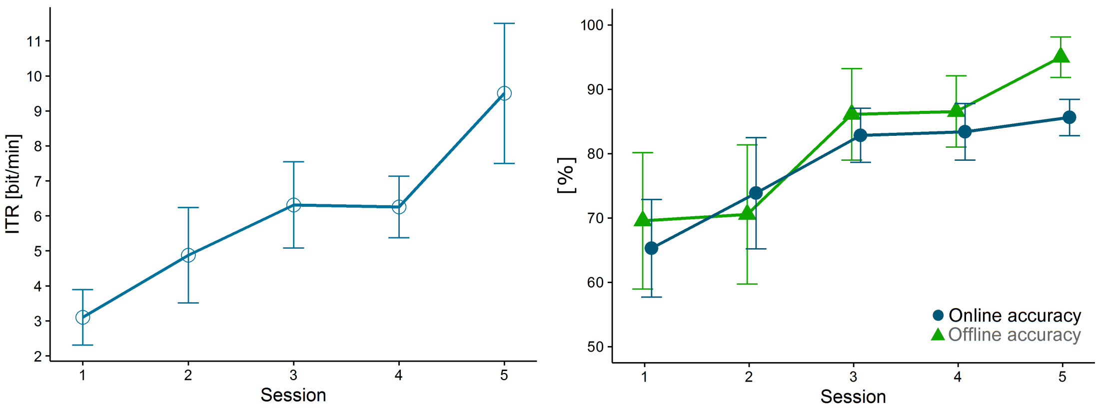

Figure 8.4: Online BCI navigation performance. Data was averaged over all BCI efficient participants (error bars represent SEM). ITR ($p = .002$), online accuracy ($p < .05$) and offline accuracy ($p < .01$) all increased significantly with training.

Table 8.2: Descriptives of BCI performance (sessions one and five) and results from Friedman ANOVAs (calculated over all sessions). Asterisks indicate significant effects.

Measure	Online Accuracy	Offline Accuracy	Online ITR [bits/min]
Session 1 (SD)	65 % (23.79)	70 % (31.86)	3.10 (2.37)
Session 5 (SD)	86 % (8.38)	95 % (9.43)	9.50 (6.01)
χ^2 (4, N = 9)	9.27	12.35	15.29
Significance (p)	.048*	.010*	.002**
Effect size (d_z)	1.00	.93	1.32

8.3.3 Effects of Age

To test for age-related differences of performance (**H4**), we split participants in the young adult and middle-aged group for this analysis. We observed a notably worse performance in the middle-aged group: Four of the six cases of BCI inefficiency occurred among the middle-aged. Since this did drastically reduce the sample size in both groups (young adult = 6; middle-aged = 3), we report descriptive ITR values only. Young adults' ITR increased from 4.49 (SD = 1.43) to 11.36 bits/min (SD = 6.53) between the first and last session, whereas the middle-aged group increased from 0.34 (SD = 0.41) to 5.79 bits/min (SD = 2.61). Thus, while online ITRs of both groups increased substantially with training, the young adults strongly outperformed the middle-aged group.

8.3.4 Questionnaires

The global average score of the NASA-TLX (calculated over all sessions and participants) was 63.3 (SD = 16.3). A Friedman ANOVA did not reveal significant changes of workload scores across sessions ($\chi^2(4,\ N = 15) = 3.22,\ p = .52$). The mean scores are summarized in Table 8.3. Descriptively, the weighted averages of the "mental demands" ($M = 20.2$) and "effort" ($M = 17.5$) indicate that these dimensions contributed most to the perceived workload, whereas "physical demands" ($M = 3.2$) and "temporal demands" ($M = 4.0$) were rated lowest. Accordingly, Bonferroni-Holm corrected Wilcoxon tests indicated that "mental demands" were rated significantly higher than "physical demands" in each session (all p <.001). Several participants proactively reported that it had been very difficult to ignore non-target vibrations, causing a high mental strain during the use of the system.

Table 8.3: Mean NASA-TLX scores across five training sessions. MD = mental demands. PD = physical demands. TD = temporal demands. P = performance. E = effort. F = frustration. Total = global NASA-TLX score. Raw = unweighted global score.

Session	MD	PD	TD	P	E	F	Total	Raw
1	17.8	3.6	2.8	13.4	19.3	9.0	60.8	50.4
2	22.5	4.0	3.4	7.5	17.7	12.0	67.3	55.5
3	21.1	4.1	3.9	8.8	18.3	12.0	67.2	55.4
4	20.5	1.7	4.9	7.8	16.0	8.6	59.5	50.8
5	19.0	2.6	4.9	7.6	16.4	10.5	61.1	52.3
Mean	20.2	3.2	4.0	9.0	17.5	10.4	63.2	52.9

Satisfaction ratings, as measured with the VAS and TUEBS, are summarized in Table 8.4. We observed a descriptive increase of the mean overall satisfaction with the BCI system (VAS) from 6.8 (SD = 2.8) to 8.0 (SD = 2.2) between the first and last session. A highly significant correlation between ITR and satisfaction ratings (averaged over all sessions) could be shown only for the full sample ($r_s = .62$, p <.01; Spearman's Rho, see Figure 8.5), but not when excluding BCI inefficient cases.

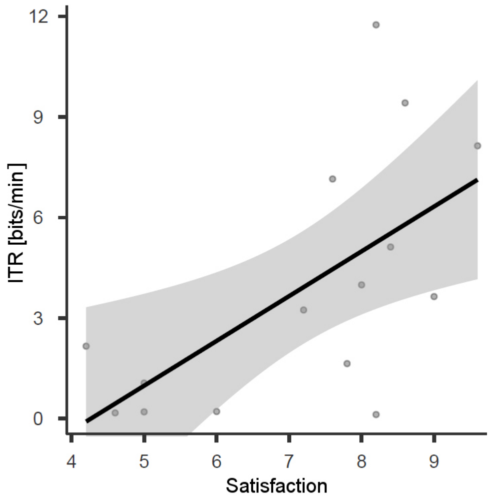

Figure 8.5: Correlation of ITR and satisfaction. Shaded area indicates the 95 % confidence interval.

The scores of the TUEBS main items remained relatively stable between first and fifth sessions ($M_1 = 4.47$ and $M_5 = 4.24$), indicating that, on average, participants felt "quite satisfied" (rating 4) to "very satisfied" (rating 5) with the BCI system. The BCI specific items received similar rating ($M_1 = 4.30$ and $M_5 = 4.18$). Paired t-tests revealed no significant effect of training for neither main or BCI specific items (both $p < .05$).

Due to the lack of evidence for significant changes of the overall scores, we report the TUEBS from the fifth session, since we assumed that participants would be more competent to judge the BCI system after the training. Furthermore, we focus on the mean scores of the BCI specific items, because we considered them to be most relevant for our study.

In the BCI part of the TUEBS, participants indicated that they felt "quite satisfied" ($M = 4.2$) with the *robustness/reliability* of the system. In a written comment, one participant speculated that electrode cables might break or that electrodes might fall off under tension or stress.

Similarly, the systems *speed* (i.e. the time required for a command) was also rated as "quite satisfactory" ($M = 3.9$). Two participants noted that it took "relatively long" or "too long" to select a command. Another participant remarked that "practice makes perfect" (quoting the German proverb "Übung macht den Meister").

The BCI's *learnability* was also rated as "quite satisfying" ($M = 4.0$) on average. One participant reported that the BCI was "effortful". Two participants (who both did not reach BCI efficiency) noted that "it probably requires a lot of practice" and that "there was no noticeable improvement", despite some "habituation".

The aesthetic design of the BCI system was rated as "more or less satisfactory" ($M = 3.2$). There were some negative remarks which mostly concerned issues with the EEG itself, specifically the cap, cables, and gel.

The last BCI specific item of the TUEBS inquired about which factors were considered most important for usability by our participants. The most frequently mentioned factors after the fifth BCI session were *ease of use* ($N = 12$), *learnability* ($N = 8$), *efficiency* ($N = 6$) and *robustness* ($N = 6$) (the TUEBS asked to select three factors).

Table 8.4: Satisfaction ratings and TUEBS scores, averaged over all participants. Parenthesis indicate ranges (values rounded to the nearest integer).

Session	Satisfaction	TUEBS (main)	TUEBS (BCI)
1	6.8 (1-10)	4.5 (3-5)	4.3 (3-5)
2	7.3 (2-10)	-	-
3	6.7 (2-10)	-	-
4	7.1 (1-9)	-	-
5	8.0 (3-10)	4.2 (3-5)	4.2 (3-5)

8.4 Discussion

This study offers an extended replication of the study by Herweg et al. (2016) with some small but notable modifications. Most importantly, we adjusted the number of sequences in the navigation task to an estimated minimum, based on each participant's individual predicted performance. Hence, we did not observe any ceiling effects as suggested by Herweg et al. (2016). Unfortunately, our sample presented with high discrepancies in the BCI performance, with many participants not reaching the threshold for BCI efficiency even after training. To still allow for a meaningful analysis, training effects were investigated among BCI efficient participants only, thus reducing our overall sample to $N = 9$, which was the same size as in the original study.

8.4.1 H1: P300 Features

A review of recent BCI studies on training effects on the P300 (see section 4.3.2) provides rather inconsistent conclusions: While some studies report amplitude increases with training (Halder et al., 2016; Herweg et al., 2016), others suggest that motivation rather than training is the most important factor (Baykara et al., 2016; Kleih et al., 2010; Nijboer et al., 2008a, 2010).

In the present study, analysis of the ERP features (amplitudes and difference between curves) did not reveal a significant effect of training, although we observed a trend toward increasing values at positions Fz and Cz. Hypotheses **H1a** and **H1b**, which predicted a training-induced increase of these physiological measures, were therefore not sufficiently supported by the present sample. Because of the small sample size, we tested with the Friedman ANOVA. However, although appropriate, this test might have been not sensitive enough to reveal minute effects and possibly resulted in one or more false negatives.

Our participants achieved peak P300 values of 6.21 μV (Cz) in the fifth session, which, despite excluding BCI inefficient cases, was considerably lower than in the study by Herweg et al. (2016), who reported values of up to 9.2 μV (Fz) after a mostly identical training.

The analysis of the P300 features in both studies was limited to data from three specific electrodes in a fixed time window. These parameters for analysis were applied to the entire group indiscriminately, possibly limiting the statistical power. An analysis of variables that are derived from a machine-learning approach, as discussed in the next section, should better account for interindividual differences, since the classifier selects discriminatory features dynamically.

8.4.2 H2: BCI Performance

The ITR from the navigation task increased more than threefold across the five sessions (3.10 to 9.50 bits/min). This observation was supported by a highly significant Friedman ANOVA (Table 8.2). We thus consider **H2** confirmed, which predicted increasing ITR with training.

This mean ITR value after training exceeds the respective 4.98 bits/min reported in the previous study by Herweg et al. (2016) almost by a factor of two. This is because Herweg et al. (2016) always used a fixed number of eight sequences for each command, regardless of participants' performances. The ITR values that can be reached were thus strictly limited. When they implemented a "bonus" task at the end of session five, in which this number of sequences was individually reduced to the number necessary to reach 70 % accuracy, they achieved a mean ITR of 20.73 bits/min.

H2 is further supported by another observation, specifically the highly significant increase of offline accuracies (Figure 8.4). Despite a lack of significant evidence for changes of the P300 features, this increase indicates that training had an effect on brain activity patterns, which in turn has led to improved classification accuracies.

Notably, this also suggests that the increase of online ITRs was not entirely caused by our reduction of the number of sequences for navigation, but also by a change in the underlying physiology. In fact, we observed an unexpected, significant increase of online accuracies with training (Figure 8.4), despite our attempt to normalize online accuracies by reducing the number of sequences according to participant performances.

Descriptively, this increase occurred mostly between the first three training sessions, after which accuracies appeared to consolidate. However, no other measure exhibited a similar pattern, rendering it difficult to interpret. It might be caused by unspecific factors only loosely related to the actual BCI system, for instance that for successful wheelchair navigation, the participants first had to become used to the novel 3D environment itself or the challenges (e.g. orientation, planning) of operating the BCI for the navigation along a course only outlined by several checkpoints.

8.4.3 H3/H4: BCI Efficiency and Age Effects

We hypothesized that after training, all participants would reach accuracies of 70 % or more (**H3**). Since only nine out of 15 participants achieved BCI efficiency, we must reject this hypothesis.

Similarly, we had hypothesized that advanced age would not have a negative effect on BCI performance (**H4**), because the results of the previous study had been rather encouraging (Herweg et al., 2016). Yet, in the present sample, the young adult group strongly outperformed the middle-aged group, so this hypothesis was not supported. Among the middle-aged group, 57 % did not reach BCI efficiency, as opposed to only 25 % of the younger group. There is little data on the prevalence of inefficiency in tactile BCI paradigms, but for visual BCIs, the percentage of inefficiency in the healthy population is estimated at only 20 % (Allison & Neuper, 2010).

This performance discrepancy between age groups could perhaps be related to the decline of sensitivity or numbers of certain mechanoreceptors, for instance Pacinian and Meissner corpuscles, as described in section 4.1 (Cauna, 1964; Iwasaki et al., 2003). The middle-aged group may therefore have been less sensitive to the tactile stimulation.

In fact, in a recent study, researchers specifically compared the EEG responses to vibrotactile stimulation in older and younger samples (Chen et al., 2019). The authors reported several significant differences between the groups, most importantly a decrease of classification accuracy in the older sample, which seems to be in line with our own observations. Chen et al. (2019) primarily attributed these results to the natural ageing of the central nervous system, but also to changes of skull thickness and decreased skin sensitivity. Additionally, an age-related decrease of attention and memory might also contribute to reduced P300 amplitudes, especially during difficult tasks (Smulders et al., 1999).

Regardless, the P300 amplitudes and the overall performances of elderly participants in the study by Herweg et al. (2016) were considerably higher than even in our young adult group. Since both studies were limited by small sample sizes, we assume that this inconsistency was caused by random sampling effects. Results on the effects of age on BCI performance thus remain inconclusive.

8.4.4 Workload, Satisfaction and Usability

The NASA-TLX was included after every session to continuously record the perceived workload during the operation of the BCI. Our analysis revealed no significant changes of the workload across the training sessions. Averaged over all sessions, the mean global score was at a relatively high value of 63.2, and thus fell into the 80 % percentile according to a recent meta study on the NASA-TLX (Grier, 2015).

As compared to similar BCI studies which also included the NASA-TLX, this score is a bit higher, albeit still in the same range: For instance, Käthner et al. (2013) reported a mean score of 57.5 from their auditory BCI. In the same study, the workload of a visual BCI was rated much lower ($M = 36.1$). However, in another example, a visual BCI received a relatively high median score of 52.3 (Riccio et al., 2015). In conclusion, there seems to be a high degree of variance even among similar paradigms. Compared to other studies, the overall workload of our tactile paradigm appears to be on the higher end of the range.

Regarding the subscores of the NASA-TLX, we found that mental demands were rated considerably higher than physical demands, suggesting that using the BCI was not physically straining, but provoked a high cognitive workload. The tactile paradigm consists of only four classes which are positioned on the body in a way that requires little or no memorization of their meaning, so it is not clear whether the stimulation can be simplified further.

Instead, the high mental demand might be an inherent property of the vibrotactile modality and its reliance on somatosensory perception, which humans are less accustomed to use, or overall less sensitive toward, as compared to the visual modality. Furthermore, it seems likely that ignoring non-target stimuli is more difficult in the tactile paradigm than for example in a visual paradigm. In the case of visual stimulation, users can control their focus on the targets by controlling their gaze, while simply not looking at non-targets (Halder et al., 2018). Such a selective perception is not possible in a tactile modality. Still, continued exposition and training in a consistent environment, might lead to an automatization of the task as described by Logan (1988) and reduce the overall mental demand.

Participant satisfaction, as measured with the VAS, ranged between 6.7 in session three to 8.0 in session five, indicating that participants were mostly satisfied with the system. This score is similar to BCI studies that also used the VAS scale to assess satisfaction: Median ratings from Riccio et al. (2015) were 6.0 - 7.2, while another comparable study reported mean ratings between 6.9 – 7.7 (Kübler et al., 2014).

We found a strong correlation between performance (ITR) and satisfaction of our participants, indicating a negative influence of performance on the satisfaction ratings. Conversely, some end-users with ALS reported high satisfaction despite low ratings of subjective performances (Holz et al., 2015b).

Finally, scores of both the TUEBS main and BCI-specific items (both at $M =$ 4.2 after session five) indicate a high general satisfaction with the BCI system. For comparison, Kübler et al. (2014) also used this version of the TUEBS to evaluate four different BCI applications, and reported similar averages, ranging from 3.7 to 4.2 for main items and from 3.5 to 4.4 for BCI-specific items. In our study, analysis of the TUEBS revealed no effect of training, possibly because five days of participation were not enough to cause substantial changes of opinion of the paradigm. Regardless, the TUEBS revealed which factors our participants considered the most important for the usability, showing that the *ease of use* and *learnability* was considered crucial.

8.4.5 Conclusion

To summarize, subjective data revealed that the paradigm provoked a high mental strain, but that participants were mostly satisfied with the system.

Regarding the effects of training, we observed a descriptive increase of physiological P300 measures and (highly) significant increases in ITR and accuracies indicating that the paradigm is trainable.

Overall, we confirmed the feasibility of this tactile BCI paradigm in a setting that resembled a potential real-life use-case, and demonstrated that efficient wheelchair control is possible, albeit not for all participants: In contrast to the previous study by Herweg et al. (2016), age may have contributed to inefficiency in our sample, but results remain inconclusive. Still, we encourage further research of this tactile paradigm with elderly participants, due to encouraging previous results.

In the present study however, not even the young adult group achieved the high performance levels reported in Herweg et al. (2016), which we attribute to random sampling effects, especially considering the small sample size of either study.

This high variance in results complicates reaching a general conclusion, but illustrates the need for designs that are tailored to the individual end-user. To facilitate this design process, it will be most valuable to identify which factors are responsible for training-related improvements of BCI performance.

9

Study II: Which Factors are Responsible for Training Effects?

9.1 Introduction and Motivation

In **study I**, an unexpectedly high number of BCI-inefficiencies had substantially reduced the size of samples available for analysis. Still, we were able to provide further evidence for training effects among participants who could use the system efficiently.

Therefore, we continued our replication efforts and recruited another, larger set of participants in order to achieve a higher statistical power. Further, to approach thesis goal 2, we wanted to explore the underlying causes of training-induced performance increases. Literature provides enough evidence to expect training effects to be related to changes in sensory perception and cognitive workload (as detailed in section 4.3.2). Thus, we put a specific focus on improvements of *somatosensory perception* and the possibility of an effect of *automatization* according to Logan (1985, 1988), and designed two additional experiments to collect evidence for both possibilities, while retaining the overall study design, which was still based on five training sessions.

Specifically, we wanted to include a condition to manipulate workload similar to the dichotic listening task of the study by Käthner et al. (2014), who had demonstrated that the hence increased workload significantly reduced accuracies of a visual BCI. During this task, participants were presented with two simultaneous stories via headphones. They had to attentively listen to one of them while ignoring the other. This observation of a negative effect of increased workload on performance was in line with several other studies which showed that a dual task condition negatively affected P300 amplitudes (Kida et al., 2004, 2012; Thurlings et al., 2013). Because of the low BCI performances in **study I** however, we decided to reduce the complexity of the listening task and expose participants to only one story at a time.

Moreover, training-induced improvements of somatosensory perception among humans have been amply demonstrated before. Nagarajan et al. (1998) used a vibrotactile interval discrimination task to assess sensitivity changes of 22 healthy participants over 10-15 days of training. Notably, the authors reported a significant change of the sensory thresholds: At the beginning of the study, participants had required a minimum relative difference of 28 % between two tactile stimuli in order to perceive them as unequal - for differences smaller than that, the two stimuli were too similar to be reliably discriminated. After the training sessions, this threshold had been reduced to only 16 %, suggesting that participants had become more sensitive.

There are relative discrimination thresholds for all sensory modalities. Interestingly, they seem to be mostly independent from the absolute stimulus differences (Kandel et al., 2000, p. 451). In fact, within the same individual and stimulus modality, the threshold applies to entirely different magnitudes: For instance, a weight difference between 100 and 120 g could just as readily be detected as a difference between 1000 and 1200 g, due to the *relative* increase of 20 % in both cases.

To determine the discrimination threshold for the somatosensory modality, Mahns et al. (2006) used a tactile frequency discrimination task with five participants and reported a threshold of approximately 30 %. Although this was a single-session study without training, the results seem to agree with the pre-training results of Nagarajan et al. (1998).

To our knowledge, a similar method to determine somatosensory sensitivity has not yet been used in the context of BCI training, but we considered it a suitable approach to assess tactile sensitivity changes, and decided to adapt it for this study.

To confirm training effects and to identify factors relevant for training, the following hypotheses were devised:

H1: As in our previous replication, we hypothesized that the physiological measures, *P300 amplitude* **(H1a)** and *difference between curves* **(H1b)**, would significantly increase across five training sessions.
H2: Further, we expected online BCI accuracies during a simple copy task (outlined in the methods below) to increase across five training sessions.

Additionally, we expected two effects regarding the possible origins of improvements between the first and last session:

H3: Firstly, we hypothesized that participants would become more sensitive to vibrotactile stimuli on the relevant body positions.
H4: Secondly, we expected that participants' accuracies **(H4a)** and P300 deflections **(H4b)** would become more robust against distraction, i.e. that any negative effect observed in session one would be decreased in session five. In line with that, we hypothesized that subjective workload would be reduced between these sessions **(H4c)**.

9.2 Methods

9.2.1 Participants

This study was conducted at the University of Würzburg, Institute of Psychology (Intervention Psychology). Healthy participants were recruited over social media and internal notice boards at the same University. Over the course of 2019, 21 participants enrolled in the study. One participant did not wish to continue after their first session because they felt not comfortable with the EEG system. Therefore, $N = 20$ participants were included (5 male, age range 20-53 years, $M = 27$, SD = 8.0).

Wanting to put a special focus on the causes of performance increases, we recruited mostly from a young adult population, as this demographic had shown better BCI performance in the previous study.

All participants were naïve with respect to BCI operation and reported normal or corrected-to-normal vision. Monetary compensation was not available for this study, but local psychology students could be compensated with course credits. All read and signed informed consent forms prior to the procedure, which was also explained verbally.

9.2.2 Procedure

Participants were invited to five sessions on separate days, with one to seven days between subsequent appointments.

At the beginning of the first and last session, we performed a tactile intensity discrimination assessment before continuing with the rest of the procedure. EEG was then recorded and processed as outlined in section 7.2. Tactile stimulation was applied using the C2 tactors as described in section 7.1. Three calibration runs were performed at the beginning of each session (section 7.3).

As opposed to **study I**, we did not include a free wheelchair navigation task for several reasons. Firstly, navigation had already been demonstrated to be feasible among healthy users twice at the time of this study. Secondly, considering our focus on training factors, we wanted to exclude any effects that were non-specific to the BCI that might occur during the navigation in a 3D environment. Most importantly however, the BCI system had no prior knowledge about the exact course which the participants intended to take. Thus, EEG data from the free navigation was not annotated with the information about the target- or non-target identity of stimuli, making it unavailable for most analyses. Instead, the linear classifier resulting from the calibration runs was used for three *copy tasks* with feedback. In the first and last session, this task was interspersed with three *dual tasks*.

At the end of each session, participants were asked to fill in a digital version of the NASA-TLX to assess workload.

9.2.3 Copy Task

Three runs of the *copy task* condition were carried out in all sessions. This task was mostly identical to the calibration runs (which comprised eight selections, for which participants had to concentrate on a specific body position announced on the screen). One notable difference was that after every selection, participants received direct feedback provided by the previously trained classifier. Secondly, the order of the eight computer-instructed target positions was pseudo-randomized (such that no direction appeared twice in a row) for each run to prevent monotony and order effects. Lastly, one selection comprised only eight sequences, as this number had been shown to be sufficient for most participants in earlier iterations of this paradigm. Unlike in the navigation task from **study I**, the number of sequences was never changed, as this design was not intended to maximize speed, but to provide controlled, comparable data for the analysis of training effects. As a consequence, the ITR as a measure of performance thus only depended on accuracy, since the time required for a selection was held constant. Although this could be regarded as a redundancy of accuracy and ITR, we report both measures to ensure comparability between studies.

9.2.4 Dual Task (Listening)

To test our hypothesis that training would lead to an increase in the robustness of BCI against distraction or concurrent activities (**H4a, H4b**), we introduced a listening task. Here, participants were instructed to closely follow a narrated story while also performing the BCI copy task. This condition was named the *dual task*.

We used two different stories depending on the session, both of which were excerpts of *Arabian Nights*, read aloud by a professional narrator, and presented over speakers. After each run, the experimenter verbally posed three questions about details of the story to possibly identify non-compliant participants in later analysis. To ensure a high level of attention, the participants were informed that they would have to answer questions about the story afterwards.

The stories in the dual task had been used previously in a study by Käthner et al. (2014), who used them in a dichotic listening task to assess the effects of mental workload on BCI performance. For the present study, this condition was simplified to only use one story at a time because the tactile BCI alone was already considered challenging. We performed a brief prestudy to confirm that attentive listening to the stories could negatively affect BCI accuracies. However, some participants' performances during early testing were still relatively robust. Because of this, the stories used in the final design were sped up by a factor of 1.25 to increase the difficulty of the task.

In the first and the last session, three runs of this dual task were performed in addition to the copy task. To control for effects of order and fatigue, copy and dual tasks were performed alternately. Conditions were balanced between participants and sessions, such that half of the participants started with the copy task in their first session, while starting with the dual task in their fifth session. The task order of the other half was inverted. The order of the two stories was balanced analogously.

9.2.5 Intensity Discrimination Task

To test the hypothesis that training would lead to an increased tactile sensitivity, we implemented a simple forced-choice intensity discrimination task, similar to approaches which analysed human tactile frequency (Imai et al., 2003; Mahns et al., 2006) or interval discrimination abilities (Nagarajan et al., 1998).

This task was performed at the beginning of the first and last session after all C2 tactors had been applied and adjusted until perceived as equally intense.

Participants received an array of short discrimination trials, each consisting of two short (220 ms each, 500 ms pause in between) vibrations applied by the C2 tactors at the right knee. The first stimulus was always set to 100 % intensity to ensure good perceptibility, whereas the intensity of second stimulus could be set by the experimenter in 5 % steps using a graphical user interface. This interface was implemented in Python 2.7 and used the tactor's own application programming interface to control the hardware. Intensities of the trials were set pseudo-randomly by the experimenter. Since a prestudy had shown that intensities below 50 % were always perceived as unequal, only intensities from 50 % to 100 % were tested.

A short auditory cue announced the beginning of a trial. To eliminate possible auditory confounders originating from the mechanical movements of the tactors, previously recorded tactor sounds were delivered via speakers such that the actual vibrations could not be heard.

After each trial, the participants were asked to verbally report whether they had subjectively experienced the two stimuli as *equally* or *not equally* strong. Repetitions were not performed if participants were unsure. Once an intensity was identified that the participant could not reliably identify as equal or unequal, this range was tested more thoroughly and extended in 5 % steps to either side until the responses were no longer ambiguous (at least 3/4 consensus of either *equal* or *unequal* responses). Depending on the participant, a different number of trials had to be performed to meet these criteria and to collect sufficient data. On average, 38 trials (SD = 6.9) per session were performed.

For analysis, the ratio of *equal/unequal* responses at a given intensity trial was projected onto a 0-1 scale, with 1 representing that only *equal* responses were given (see Figure 9.1). The discrimination threshold was defined as the intensity value at which *equal* and *unequal* responses were given with the same probability (0.5). This point was estimated automatically by fitting the data with a sigmoidal function using a nonlinear least squares method. In four out of 20 cases, a sigmoidal fit was not possible, because ratios jumped almost immediately from 0 to 1. Those distinct thresholds were estimated manually (shown in supplementary Figure A.3). Using this method, the stimulus intensities at 0.5 were extracted for all participants.

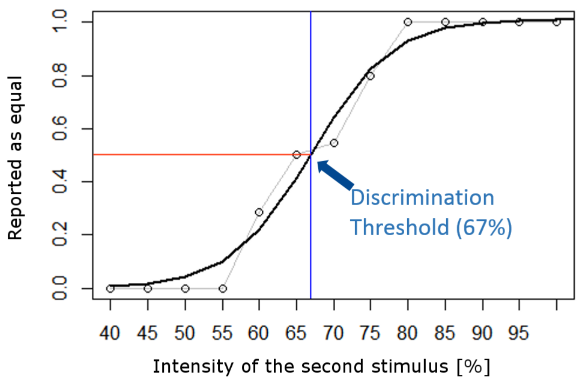

Figure 9.1: Data from one intensity discrimination task. Points indicate the ratio of *equal/unequal* responses at the respective intensities. In the range between 60 % and 75 %, responses were ambiguous. The discrimination threshold, at which both answers would be given with the same probability (0.5), was estimated using a sigmoidal fit.

9.2.6 Statistical Analysis

Since calibration and copy tasks were designed to be very similar, EEG data from both was combined for analysis to benefit from the hence enlarged data size and improved signal-to-noise ratio. EEG data preprocessing was performed as outlined in section 7.2, using a rejection threshold of ± 100 μV and a time window of 300-500 ms. As before in **study I** and Herweg et al. (2016), we focused on positions Fz, Cz and Pz for this analysis.

For every participant and session, mean amplitudes and mean differences were averaged over all calibration and copy runs of a session. Extracted amplitudes and differences between curves entered two separate repeated measures ANOVA with the factors *session* (five levels) and *electrode* (three levels). The average online accuracy of the three copy runs entered a repeated measures ANOVA with the factor *session* (five levels). The same was done for online ITRs. Effects were further analysed with planned t-test between first and last sessions.

The sensory thresholds from the discrimination task on sessions one and five entered a repeated measures t-test to detect training effects.

For analysis of the dual task, the impact on performance was evaluated by comparing the accuracy difference between dual and copy task conditions (sessions one and five only).

9.3 Results

Twenty participants successfully completed all five BCI training sessions, which resulted in a total of 100 full BCI sessions available for analysis. Nineteen participants reached BCI efficiency (using the same criterion as before, i.e. accuracy $\geq$ 70 % during the last, or at least three other sessions). Participant 3 did not achieve this threshold and was not included in further analysis or grand average plots.

9.3.1 Physiological Measures

Figure 9.2 shows the grand average of all BCI efficient participants at positions Fz, Cz and Pz. Data from session one and five are plotted to visualize the development with training. Descriptively, mean amplitudes and differences between curves increased in the P300 range at Fz and Cz. Highest P300 peak values (in the P300 range) were observed at 5.05 μV after 408 ms (Fz) and at 5.20 μV after 369 ms (Cz) in session five.

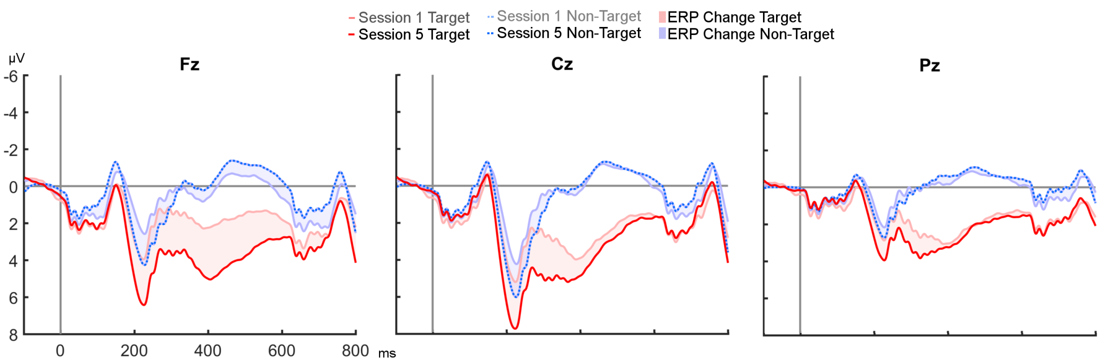

Figure 9.2: Grand averages from the first and last sessions from the central midline. Shaded areas indicate changes between first and last session.

Figure 9.3 shows the development of the extracted mean amplitudes and mean differences across the five training sessions. Descriptively, the largest increase occurred between the first and second session at Fz. Position Cz started out from higher values, but the largest increase was also observed between the first and second session. Descriptively, no or only minor increases were observed after session four at any electrode position.

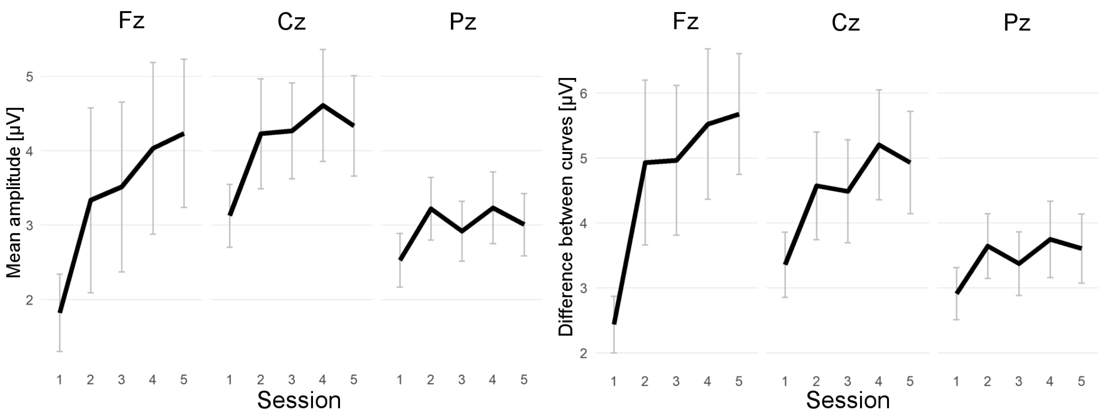

Figure 9.3: ERP features across all sessions (error bars represent SEM).

A repeated measures ANOVA of amplitudes revealed no main effect of *electrode*, but a main effect of *session* and an interaction effect of *session* and *electrode*. Due to violations of the assumption of sphericity, a Greenhouse-Geisser correction was performed, after which the main effect of session only reached marginal significance, but a high effect size ($F_{(1.8,\ 32.4)} = 3.021$, $p = .067$, $\eta p^2 = .144$). The interaction effect was not significant ($p = .113$) after correction.

Similarly, the repeated measures ANOVA of the difference of the curves revealed no main effect of *electrode*, but a main effect of *session* and an interaction effect of *session* and *electrode*. After Greenhouse-Geisser corrections, the main effect of *session* was not significant ($F_{(2.05,\ 36.97)} = 2.639$, $p = .084$, $\eta p^2 = .128$). The interaction effect however, remained highly significant, with a high effect size ($F_{(3.07,\ 55.26)} = 4.808$, $p < .01$, $\eta p^2 = .211$).

Figure 9.4 shows the average target and non-target epochs at Fz from session one and five for all participants individually (Fz was chosen since training effects were strongest at this position; for Cz and Pz, see supplementary Figures A.4 and A.5). Descriptively, there was a large diversity of the ERPs and the effect of training on the individual level: In some participants, the P300 deflection increased considerably with training (Participants 6, 8, 10, 14, 19). Others did not show a strong divergence of the curves, but achieved efficiency regardless. Participants 4 and 17 appeared to exhibit an inverted polarity of the putative P300.

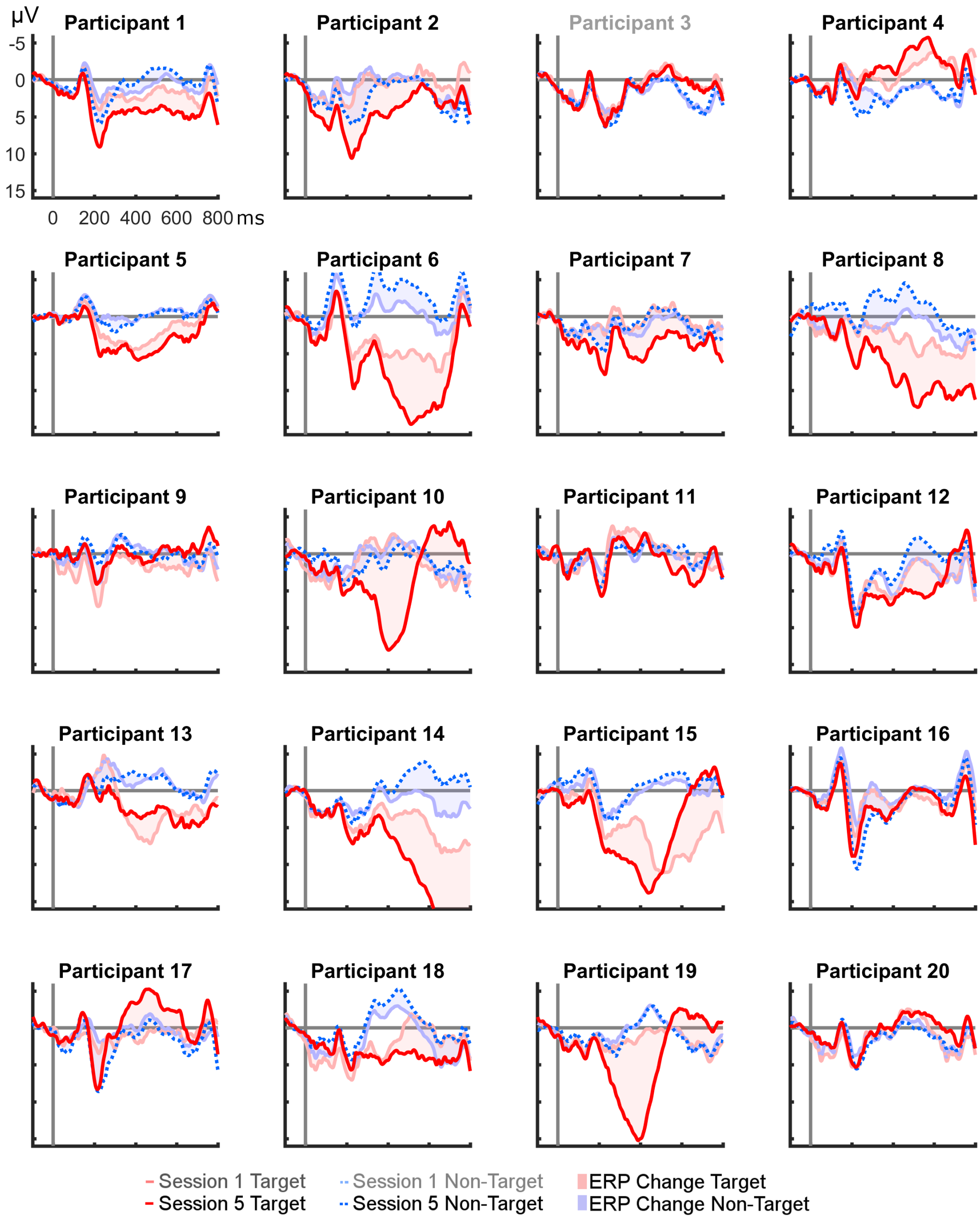

Figure 9.4: Individually averaged post-stimulus epochs at Fz of all participants' first and last sessions. Shaded areas indicate changes. Inefficiencies are indicated by a gray participant code.

To further investigate this interaction, we calculated one-tailed paired t-tests for electrodes Fz, Cz and Pz between the first and last session, where we expected the strongest effect. Driven by our hypotheses for a training effect between the sessions (**H1a, H1b**), this was done for both amplitudes and differences. After Bonferroni correcting for multiple electrode comparisons, we found a significant amplitude increase on Fz and significant increases of the difference between curves at Fz and Cz. Table 9.1 provides a summary of the descriptives and t-test results and effect sizes (Cohen's d_z).

Table 9.1: Overview of the descriptives of amplitudes and differences between curves of session one and five. Asterisks indicate significant differences between sessions one and five.

Electrode Position	Fz	Cz	Pz
Amplitude S1 [μV]	1.82 (SD 2.26)	3.13 (SD 1.84)	2.53 (SD 1.57)
Amplitude S5 [μV]	4.23 (SD 4.34)	4.33 (SD 2.94)	3.01 (SD 1.82)
t-test (p)	.023*	.075	.330
Effect size (d_z)	.61	.49	.29
Difference S1 [μV]	2.44 (SD 1.90)	3.36 (SD 2.19)	2.91 (SD 1.75)
Difference S5 [μV]	5.68 (SD 4.05)	4.93 (SD 3.44)	3.61 (SD 2.32)
t-test (p)	.002**	.014*	.125
Effect size (d_z)	.93	.67	.42

9.3.2 BCI Performance

During the last session, participants 2, 5, 6, 14, 16 and 17 completed all three copy tasks without any mistakes, whereas only one participant (18) had achieved the same in session one.

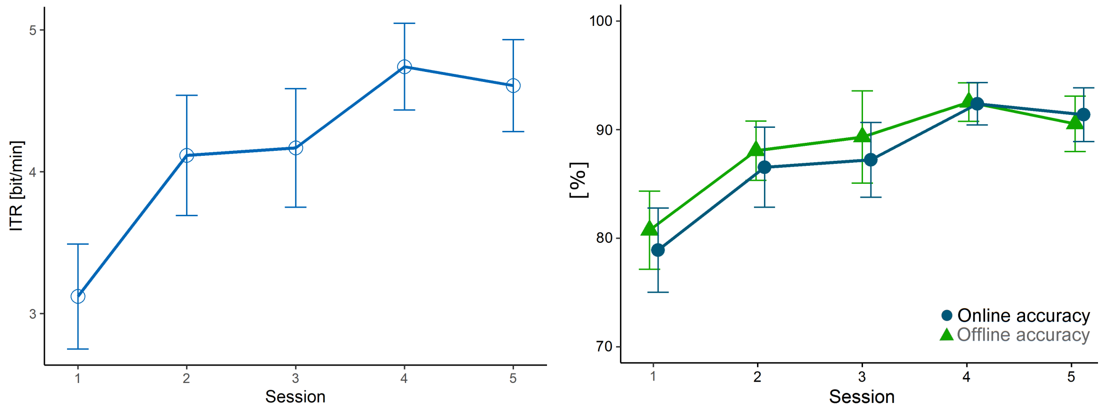

Figure 9.5: Online BCI copy task performance. Data were averaged over all BCI efficient participants (error bars represent SEM). All increased significantly with training.

The development of online/offline accuracies and online ITRs from the BCI efficient participants across the five training sessions is visualized in Figure 9.5. Descriptively, the largest performance increases were observed between sessions one and two, whereas a slight decrease occurred between the fourth and fifth session.

To test for training effects on these measures, we calculated three repeated measures ANOVAs with the factor *session* (5 levels). The tests revealed a highly significant main effect of training session ($F_{(4,\ 72)} = 4.351$, $p = .003$, $\eta\text{p}^2 = .195$) on the online accuracies from the copy task. Between the first and last sessions, mean online accuracies increased from 78.7 % to 91.2 % (t-test $p = .009$). These values exceeded the approximate chance level threshold of 50 % (alpha = .01; Combrisson & Jerbi (2015); Müller-Putz et al. (2008)).

Similarly, a main effect of session was revealed on the cross-validated offline accuracies (calculated using all 10 sequences from the calibration data) ($F_{(4,\ 72)} = 2.959$, $p = .025$, $\eta p^2 = .141$). Values increased significantly from 80.74 to 90.53 % between the first and last session (t-test, $p = .038$).

Further, online ITRs increased highly significantly ($F_{(4,\ 72)} = 4.847$, $p = .002$, $\eta p^2 = .212$) from a mean of 3.12 to 4.61 bits/min between first and last sessions (t-test, $p = .004$). Table 9.2 provides an overview of the descriptives and the results from the contrast between session one and five.

Table 9.2: Descriptives of BCI performance and planned t-tests results between sessions one and five. Asterisks indicate significant effects.

Measure	Online Accuracy	Offline Accuracy	Online ITR [bits/min]
Session 1 (SD)	78.7 % (16.8)	80.74 % (15.67)	3.12 (1.61)
Session 5 (SD)	91.2 % (10.7)	90.53 % (11.12)	4.61 (1.41)
t-test (p)	.009**	.038*	.004**
Effect size (d_z)	.67	.51	.76

9.3.3 Dual Task Effects

After listening to the story in session one and five, participants were able to answer M_1 = 7.05, (SD = 1.82) and M_5 = 7.8 (SD = 1.24) out of nine questions correctly. The lowest and highest scores overall were 3 and 9. We did not exclude participants with low question scores, since their dual task accuracies were decreased nonetheless.

On a behavioral level, we found a negative effect on accuracies during the dual task condition. Table 9.3 provides an overview of the number of correct classifications for all participants. In the first session, copy task accuracies (M = 78.9 %, SD = 16.9) were significantly higher than in the dual task condition (M = 70.9 %, SD = 21.5), as revealed by a paired t-test (p = .038). In session five, both copy (M = 91.4 %, SD = 10.8) and dual task (M = 81.3 % SD = 16.5) accuracies ranged on an overall higher level, but there was an even larger difference between the tasks (p = .001).

To directly test for between-sessions effects on the impact of the dual task condition on online accuracies, we calculated the *relative* performance differences (Δ_{rel}). To achieve this, the difference between the conditions (*dual-copy*) was expressed as a fraction of the copy task value ($\Delta_{\text{rel}} = \frac{dual-copy}{copy} * 100$). However, a paired t-test (session 1 vs. 5) of this relative difference revealed no significant result (p = .835).

Table 9.3: Number of correct selections in the *dual* and *copy* tasks (out of a maximum of 24). The *difference* indicates the negative impact of the dual task.

Participant	1	2	3$^{\text{ex}}$	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	M
S1 Copy	21	20	14	21	14	22	10	17	13	21	22	20	21	22	21	23	12	24	19	16	18.9
S1 Dual	18	17	7	20	16	22	4	14	6	18	15	20	17	21	22	19	20	21	22	11	17.0
Abs. Difference	-3	-3	-7	-1	2	0	-6	-3	-7	-3	-7	0	-4	-1	1	-4	8	-3	3	-5	-1.9
Rel. Difference [%]	-14	-15	-50	-5	14	0	-60	-18	-54	-14	-32	0	-19	-5	5	-17	67	-13	16	-31	-10.3
S5 Copy	23	24	12	22	24	24	19	21	22	23	23	21	14	24	23	24	24	21	22	18	21.9
S5 Dual	21	22	18	22	23	24	14	17	13	22	18	18	10	23	22	23	18	23	20	17	19.5
Abs. Difference	-2	-2	6	0	-1	0	-5	-4	-9	-1	-5	-3	-4	-1	-1	-1	-6	2	-2	-1	-2.4
Rel. Difference [%]	-9	-8	50	0	-4	0	-26	-19	-41	-4	-22	-14	-29	-4	-4	-4	-25	10	-9	-6	-11.5

Notes: M = Mean over all included participants. ^{ex} *= excluded.*

On the physiological level, visual analysis indicated minor differences in the shape of the ERPs, with the dual task showing reduced amplitude spikes at 180 and 220 ms (Figure 9.6). In session one, mean target amplitudes appear to be larger in the dual task. Other than that, no major differences of the conditions could be observed.

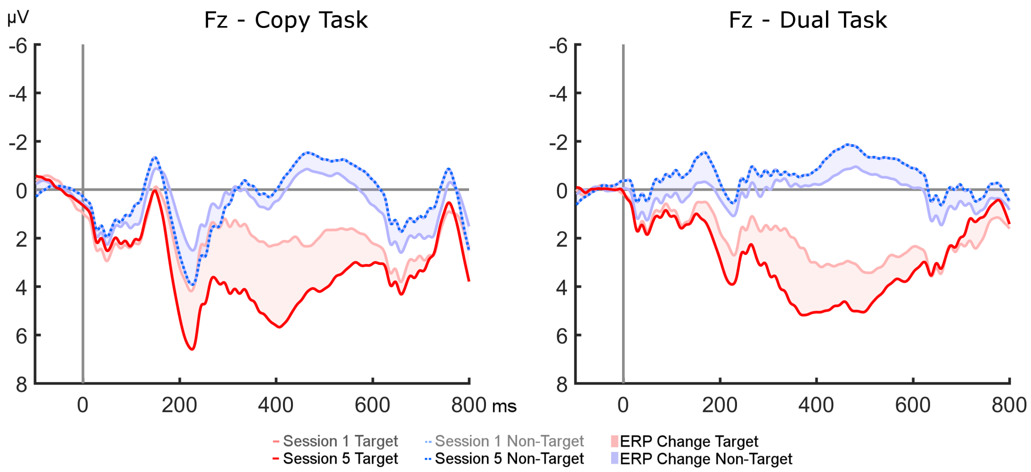

Figure 9.6: Deflections elicited during the copy and dual task conditions in sessions one and five at electrode position Fz.

For statistical analysis, we focused on the difference between curves at position Fz, since we had observed the strongest amplitudes and training effects here (see Table 9.1). Paired t-tests revealed a significant effect of task on the difference between curves in session one (p = .050, d_z = .47). Values from the dual task were higher than from the copy task (M_{dual} = 3.94 μV, M_{copy} = 2.70 μV). In the fifth session however, this effect disappeared (p = .596, d_z = .12), and the values were closer together (M_{dual} = 7.30 μV, M_{copy} = 6.59 μV).

Again, we calculated the relative difference between the values of the copy and dual tasks to test for an effect of training on the impact of the distraction, as described above. The average relative difference between copy and dual task values decreased from M_1 = 44,4 % to M_5 = 12.8 %. However, a paired t-test (session 1 vs. 5) revealed no significant result (p = .465).

9.3.4 Vibrotactile Sensitivity

A pre/post-training comparison of vibrotactile intensity discrimination sensitivity revealed a highly significant decrease of the stimuli difference necessary to be perceived as unequal (t-test $p < .001$, $d_z = .99$). During the first session, the second stimulus had to be reduced by an average of 25.6 % (SD = 7.8) for participants to reach the defined discrimination threshold (i.e. the point of equiprobable "equal" and "unequal" responses). In the last session, intensities had to be reduced by an average of 20.7 % (SD = 7.8). Thus, most participants were able do discriminate between smaller intensities after training (see Figure 9.7).

Further, we explored the correlation between the thresholds and BCI accuracies (copy task). However, no significant Pearson correlation on the first ($p = .575$, $r_p =$ -.137) or the last session ($p = .885$, $r_p =$ -.036) was found (Figure 9.8).

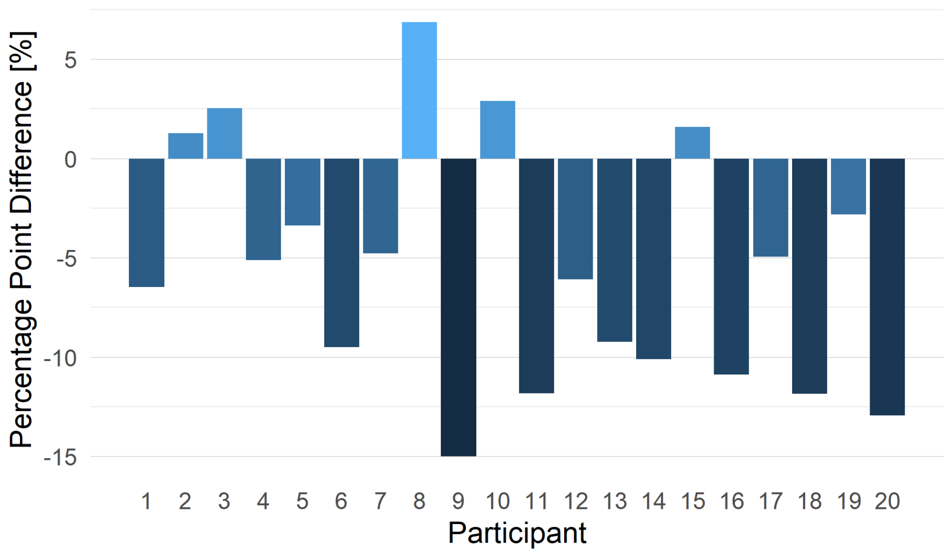

Figure 9.7: Change of the discrimination thresholds between first and last session. Negative values indicate a decrease in the threshold, meaning that smaller intensity differences could be discriminated (for instance, participant 1 initially required a 34 % stimulus difference, but only 28 % after training). BCI inefficient participant 3 is shown but not included in statistical analysis.

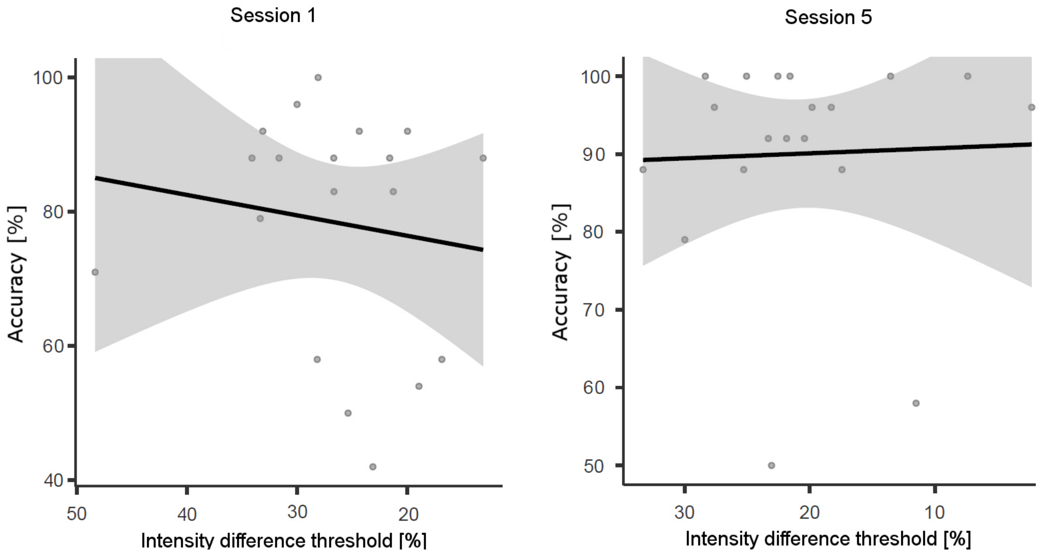

Figure 9.8: Correlations of the sensory threshold and accuracy (both n.s.). Low threshold values indicate a high discriminatory ability. Shaded areas show the 95 % confidence interval.

9.3.5 Workload

The global average score of the NASA-TLX (calculated over all sessions and participants) was 55.8 (SD = 21.1). A repeated measures ANOVA revealed a significant effect of session on the total workload scores across sessions ($F_{(4,\ 72)} = 3.402$, $p =$ .013, ηp^2 = .159). Post-hoc comparisons revealed that the workloads from middle sessions (two to four) were significantly lower ($p < .05$) as compared to sessions one and five, which included the sensitivity assessment and the dual task (see Figure 9.9). No further significant workload differences were found, however, neither between first and last or between middle sessions. Workload continuously decreased between session one and four. Thus, with the exception of session five, there was a descriptive trend toward lower workload scores with training. All mean scores are summarized in Table 9.4.

The weighted averages of the subscores "mental demands" (M = 18.2) and "effort" (M = 15.5) indicate that these dimensions contributed most to total workload scores, whereas "physical demands" (M = 1.1) were rated lowest. Bonferroni adjusted t-tests indicated that "mental demands" were rated significantly higher than "physical demands" in all sessions (all $p < .001$).

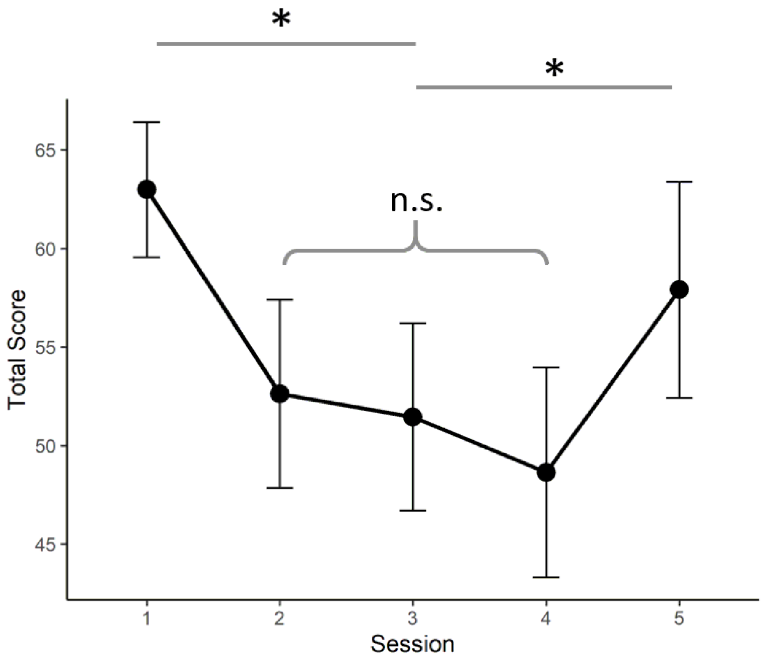

Figure 9.9: NASA-TLX total scores. Despite a descriptive trend toward lower workload with training, first and last sessions were scored significantly higher than the rest.

We were interested in the "mental demands" subscore because of its relevance regarding the dual task condition, and calculated a paired t-test between the first and fifth session, which revealed a highly significant decrease from an average of 23.7 (SD = 6.9) to 17.4 (SD = 9.8) (p = .002).

To test whether the mental demands, as measured by the NASA-TLX, correlated with the expected negative impact on accuracies of the dual task, we calculated the *relative* accuracy differences between the conditions for each participant (as described above on page 91). A Pearson test, however, revealed no significant correlation in the first (p = .423, r_p = .208), or the last session (p = .348, r_p = .243) (see Figure 9.10).

Table 9.4: Mean NASA-TLX scores across five training sessions. MD = mental demands. PD = physical demands. TD = temporal demands. P = performance. E = effort. F = frustration. Total = global NASA-TLX score. Raw = unweighted global score.

Session	MD	PD	TD	P	E	F	Total	Raw
1	23.7	0.1	6.1	7.1	18.5	8.0	63.6	50.9
2	18.3	0.5	5.5	6.7	14.8	7.7	53.6	43.5
3	16.7	1.5	7.9	6.3	13.8	6.8	52.9	42.5
4	14.5	1.7	6.6	6.3	15.3	5.3	49.8	39.7
5	17.4	1.9	7.7	8.8	15.2	7.8	58.9	47.1
Mean	18.2	1.1	6.8	7.0	15.5	7.1	55.8	44.7

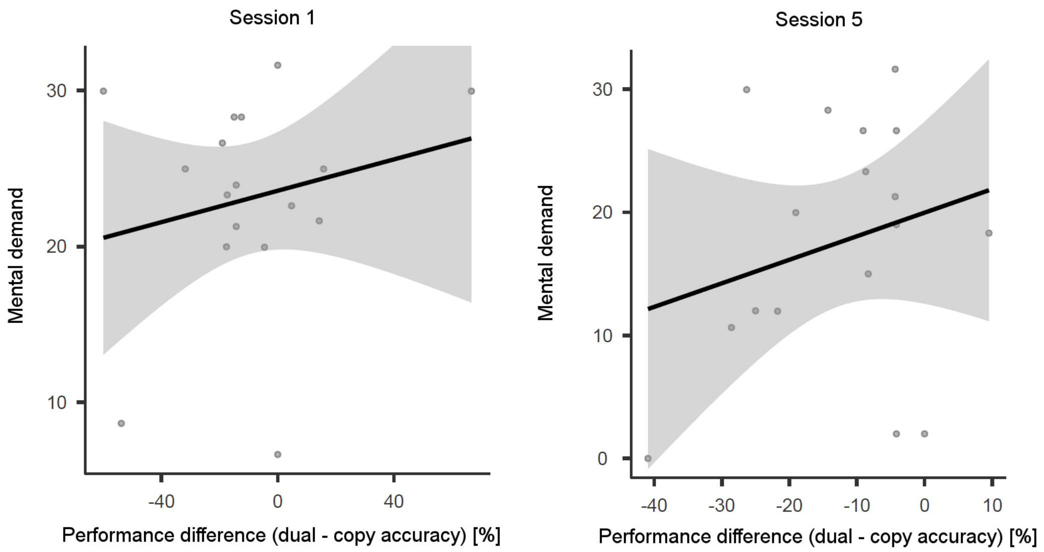

Figure 9.10: Correlations between *mental demands* and the relative accuracy reduction during the dual task (both n.s.). A positive performance difference indicates that accuracies in the dual task were higher than in the copy task. Shaded areas indicate the 95 % confidence interval.

9.4 Discussion

The present study provides further evidence for training effects with this tactile paradigm. As compared to Herweg et al. (2016) and **study I**, results were based on a substantially larger sample, which was not affected by a high number of BCI inefficiencies. Most importantly though, we offer first clues as to which factors may account for training-induced performance increases of tactile BCI paradigms. To identify these factors, additional tasks were included in the first and last session, increasing the overall workload. This may explain why ERP and performance measures appeared to strongly increase after the first session, but to plateau after the fourth.

9.4.1 H1: P300 Features

Despite strong descriptive increases in the mean P300 amplitudes across the five sessions, the omnibus test (ANOVA) revealed no consistent main or interaction effects, and thus provided no conclusive support for hypothesis **H1a**. This may be due to a high variance between participants, who in some cases showed no observable deflections or deflections of an opposite polarity. However, a hypothesis-driven contrast of session one and five revealed a significant amplitude increase at electrode Fz. Observed amplitudes at Cz were higher overall (both before and after training), but did not increase as much across sessions.

The ANOVA of the difference between curves revealed no main effects, but a significant interaction effect (*session* * *electrode*), indicating that there was a significant effect of training, but not for all electrodes. Testing individual electrodes then revealed training effects for Fz and Cz, supporting **H1b**.

Taken together, these results provided sufficient evidence to cautiously accept the general hypothesis **H1** of a training effect on physiological measures.

The fact that evidence from the difference between curves was stronger than just the target amplitudes suggests that there has been an effect of training on the non-target responses. Comparing session one and five, visual analysis revealed a notable shift of the non-target curves into the negative polarity, increasing the difference between curves (Figure 9.2, Fz). This has also been observed descriptively in **study I** (Figure

8.2, Fz) and to some degree in Herweg et al. (2016). This might be due to a contamination of the non-target epochs with unintended P300-like deflections in response to the sometimes startling, highly salient nature of the vibrotactile stimuli. Although this remains speculative, ignoring non-target stimuli can be very hard in the tactile modality, as proposed before in **study I**, section 8.4.4. By the fifth session, participants may have learned to better ignore non-target stimuli, explaining how the non-target epochs became more negative over time.

While this does not fit within the classic model of the oddball paradigm, there are several studies that describe that obtrusive (auditory) stimuli can automatically evoke a P3b-like response even when attention is deliberately directed away (Polich & Kok, 1995; Putnam & Roth, 1990; Roth et al., 1982, 1984). Furthermore, in a study using visual and auditory three-stimulus paradigm, Comerchero & Polich (1999) found that non-target amplitudes were larger (i.e. more positive) when target/non-target discrimination was difficult. Specifically, they reported an enlarged P3a deflection at frontal electrode sites, which could explain the positive shift of the non-target curves we observed at Fz (Figure 9.2) in the first session.

In comparison, mean amplitudes of 4.33 μV were slightly higher than the 3.94 μV measured in **study I** (both at Cz). The highest observed peak was at 6.21 at μV (Cz), which was still considerably lower than the 9.2 μV (Fz) reported by Herweg et al. (2016).

9.4.2 H2: BCI Performance

All measures of BCI performance (online and offline accuracy, online ITR) showed significant and substantial increases with training. Hence, our respective hypothesis (**H2**) could be confirmed without reservation.

Mean accuracies of 91.2 % in session five were not only significantly over the chance level (Müller-Putz et al., 2008), but also nearly as high as the 92.56 % reported by Herweg et al. (2016), making this the second highest mean accuracy of comparable studies summarized in the general introduction in Table 5.2. Similarly, offline accuracies in session five reached a mean of 90.53 %, a value which is consistent with the online accuracies.

Furthermore, our mean online ITR from session five of 4.61 bits/min was very close to that of Herweg et al. (2016), who reported 4.98 bits/min, while also using a fixed number of eight stimulations per selection. In the present study, we did not adapt this number based on the participants performance, as our focus was on aspects other than speed. The high values of **study I** and the bonus task from Herweg et al. (2016) were therefore not achievable.

In a five-session training study of an auditory paradigm, Baykara et al. (2016) reported that performance increases saturated after session three. We found no further performance increase between session four and five, but this may be due to the two additional task that increased the duration and effort of session five.

9.4.3 H3: Vibrotactile Sensitivity

The stimulus difference at threshold levels decreased considerably from a mean of 25.6 % to 20.7 % between the first and last sessions, indicating that training enabled participants to perceive smaller stimulus discrepancies. This highly significant result, together with a large observed effect size, shows that the vibrotactile sensitivity of our participants has increased across the training sessions. We thus consider **H3** confirmed.

Our finding is in line with numerous other studies which reported training effects on tactile sensitivity (Imai et al., 2003; Nagarajan et al., 1998; Reuter et al., 2014). These studies used slightly different approaches and used temporal, spacial or frequency discrimination, so a quantitative comparison with the results of our intensity discrimination is difficult. Instead, we provide strong, novel evidence for training effects for a largely undescribed dimension of tactile sensitivity improvement.

This effect of training on sensitivity suggests that sensory perception may in fact be an important factor for tactile BCI improvements. Although we could not demonstrate a correlation between BCI performance and sensitivity for either session, literature provides evidence that high stimulus discriminability facilitates P300 elicitation (Comerchero & Polich, 1999; Waal et al., 2012). Hence, while the full implications of these first results have yet to be explored, it seems plausible that increased sensitivity contributed to BCI performance by facilitating stimulus discrimination. What remains unknown is where exactly along the somatosensory pathway this adaptation may have taken place.

It should be noted that the highly significant training effect was found despite certain confounding factors, for instance the fact that tactors were attached over the clothing. Since we did not instruct the participants in regard to their clothing choices, imprecise measurements might have contributed to an amount of noise in the data. To counteract this, we did not determine absolute perception thresholds, but relative stimulus differences between two stimuli which were both strong enough to be perceived regardless of clothing.

Furthermore, differences in weather conditions between first and last sessions may have affected the choice of outfit and thus the thickness of the fabric between the tactors and the skin, but it is unlikely that this introduced a bias in any specific direction. This study was conducted over several seasons, which should have evened out any confounding influence of weather-related clothing choices. Additionally, participation in all five sessions only took about 1.5 weeks, and thus happened largely within the same yearly season.

9.4.4 H4: Workload and Dual Task Effects

General Workload and Training Effects

The principal results from the NASA-TLX were similar to those of **study I**. Again, the discrepancy between mental and physical demands confirmed that the BCI was not physically straining, but provoked a high cognitive workload. The mean global score of 55.8 was lower than in **study I** (63.2), falling into the 60 % percentile according to the meta study on the NASA-TLX (Grier, 2015). In fact, the score was closer to those reported in the BCI studies discussed earlier (section 8.4.4; Käthner et al., 2013; Riccio et al., 2015). It seems likely that the high number of BCI inefficiencies in **study I** have caused the comparatively high subjective workload, whereas the sample of the present **study II** was not affected in such a way.

Moreover, it appears that the addition of the dual task, which was a focus of the study, had increased the subjective workload in sessions one and five. Hence, while the total scores continuously decreased until session four, they went up to their second highest value in session five. While this was not surprising, it complicated comparisons across all sessions.

Comparing the descriptive decrease between first and last sessions, which were comparable due to their identical design, hypothesis **H4c** (decreasing total workload) was not clearly supported, as the test revealed only marginal significance. Yet, a significant reduction of the "mental workload" subscore supports the hypothesis, albeit only for this dimension. This especially might be a manifestation of an automatization of the task as described by Logan (1988), but the training program may not have been sufficient enough to cause demonstrable effects on the overall workload.

Dual Task

Judging by the significant negative impact of the dual task on online accuracies, we assume that the manipulation was successful.

Hypothesis **H4a** (regarding a convergence of performances) however, could not be supported, as copy- and dual task accuracies appeared to diverge more, not less, with training. On the physiological level, we found a significant, task-related difference in the P300 deflections in session one, but not session five. Despite this observation, we could not confirm our hypothesis that the P300 would be less affected by a dual task after training **(H4b)**, since the difference in session one was actually caused by the dual task amplitudes being *larger*, not smaller as we had predicted (which suggests that with training, dual task deflections actually became affected more, not less). A cause for this could not be identified based on the present data.

Thus, while we confirm previous observations that a dual task condition similar to Käthner et al. (2014) negatively affects BCI performance (Thurlings et al., 2013), we could not demonstrate this on the physiological level.

Despite testing the dual task's effect on performance in a brief pre-study, we may have underestimated the paradigm's robustness against such manipulation, possibly due to the poor BCI performances from the previous study. Hence, the distracting effect of this task may not have been strong enough to become apparent in the analyses of training effects and physiological measures. Furthermore, a relationship between the dual task impact on accuracies and the "mental demand" scores from the NASA-TLX could not be shown.

In conclusion, although we confirmed the negative effect of the dual task (i.e. increasing overall workload decreases performance), none of our **H4**-hypotheses were sufficiently supported by the present sample. Thus, besides a notable decrease in mental workload overall, we found no evidence for workload as a factor for training-related BCI performance increases.

9.4.5 Conclusion

With this study, we provided further strong evidence not only for the feasibility, but also trainability of the tactile paradigm, backed by relatively large sample size.

In comparison to Herweg et al. (2016), the performance levels reported in the present study were very similar, and both were substantially higher than those of comparable studies (Table 5.2). This confirmed that the present paradigm is currently one of the best options for BCI based on the tactile modality, and is perhaps the most promising approach to establish among impaired end-users.

Regarding our endeavour to identify possible factors for training, we unexpectedly found no evidence to support a training-related increase of the BCI's robustness against distraction as induced by a dual task condition. On the other hand, we demonstrated that the paradigm was relatively stable and performed significantly above chance level even when listening attentively to a human voice. This suggests that a potential real life application, in which distraction is sometimes inevitable, might not be overly impeded either.

Finally, we found evidence that performance increase may be mediated by improvements along the somatosensory pathways (section 4.1). Thus, further investigation into this aspect seems merited.

10

Study III: Tactile BCI Operation by ALS Patients

10.1 Introduction and Motivation

Although basic research on healthy participants is a crucial first step, the transferability of these results to impaired patients may be rather limited. With the general feasibility of the tactile paradigm established, it is now time to take a closer look at potential end-users.

Unfortunately, the translational gap between research in the laboratory and actual home-use by patients in need of assistance is still substantial (Kübler, 2013; Kübler et al., 2015; Nijboer, 2015). While there are a number of studies which investigated BCI operation by impaired users (e.g. Kübler et al., 2005; Nijboer et al., 2008b; Sellers & Donchin, 2006) there are very few publications that provide long term data from patients that allow to investigate effects of training and disease progression (e.g. Botrel et al., 2017; Kübler et al., 1999; Sellers et al., 2010).

Notably, Sellers et al. (2010) found that over a period of 2.5 years of independent home use, neither BCI performance nor P300 amplitudes had deteriorated. Conversely, Holz et al. (2015a) performed a long-term study on two ALS patients, who independently used Brain Painting over a period of several years. The authors reported a decrease of the P300 amplitudes in one of the participants and suggested ALS-related neuronal degeneration, described in several other studies, as a possible reason (Hanagasi et al., 2002; Paulus et al., 2002). Amplitudes from the second patient remained mostly stable. The study also implemented the UCD approach, collecting subjective data from several questionnaire items.

Study III was specifically designed to evaluate the tactile BCI among potential end-users, so that in the future, the paradigm could be better tailored to the individual in need. Thus, we hope to contribute to closing the translational gap and that the end-users may be able to benefit from the technology soon.

Specifically, we wanted to approach thesis goal 3 and establish the paradigm's feasibility among ALS users and explore possible long term effects. To this end, we established a cooperation with the University of Aalborg, Denmark, where several ALS patients were locally recruited to participate in a *cross-sectional* study, and with the Technical University of Nürnberg (TH Nürnberg), which provided engineering and practical support for a *longitudinal, single-case* study with one ALS patient.

While the cross-sectional (single-session) study was exploratory, we formulated several hypotheses regarding training effects and usability for the single-case study:

H1: Firstly, we expected an increase of the physiological measures of the P300 with training. This was supported by Silvoni et al. (2016), who found no pathology of tactile ERPs in ALS patients.
H2: BCI performance was predicted to increase across training sessions, in line with previous findings from Herweg et al. (2016) and **studies I** and **II**.
H3: Secondly, we hypothesized that training would lead to a decrease both in overall workload (**H3a**) and specifically in the mental demand (**H3b**) of the task (Logan, 1988).
H4: Finally, we expected user satisfaction to correlate with BCI accuracy (**H4a**) and to subjective control (**H4b**), similar to Holz et al. (2015a).

10.2 Methods

10.2.1 Participants

Cross-sectional (Aalborg)

Seven participants with various stages of ALS were recruited by the University of Aalborg. Table 10.1 provides an overview of their ALS history and condition at the time of participation. Data collection, including the assessment of speech and the *revised ALS functional rating score* (ALSFRS-R) was performed by scientific and medical staff of the University of Aalborg. The ALSFRS-R is a 12-item questionnaire designed to assess the patient's self-sufficiency (Brooks et al., 1996; Cedarbaum et al., 1999). All items range from 0-4, with 4 indicating full self-sufficiency. Thus, a maximum total score of 48 could be reached. The study was performed at the laboratory of the University of Aalborg.

Table 10.1: Overview of patients.

Patient code	Sex	Age [y]	Years since diagnosis	Years since symptom onset	Speech	ALSFR-R
02	M	62	1.67	2.5	slightly impaired	42
05	M	78	7	8	impaired	36
06	M	53	0.67	1.5	normal	39
07	F	58	0.67	1.35	no speech	25
08	M	73	0.75	1.5	normal	47

Note: For ALS patients 11 and 12, no detailed medical history data was available.

Longitudinal (Nürnberg)

The patient from Nürnberg was born in July 1962. He is an alumni of the TH Nürnberg with a degree in engineering. After receiving the first diagnosis of ALS in 2010, he has been in a state of the *incomplete* LIS (Bauer et al., 1979) for about three years at the beginning of the study, leaving him only blinking and eye movements, with which he communicates using simple yes/no responses, morse code, partner scanning, and with which he operates his personal computer using eye tracking software. His ALSFR-R score was 0 at the beginning of data collection in February 2018. He is artificially ventilated continuously, fed via PEG and medicated with *Riluzole*. His vision is corrected-to-normal.

The patient participated in our study in his own home with the help of his caretakers and an experimenter from the TH Nürnberg. In the following, this patient will be abbreviated as *PT-N* to avoid confusion with the numeric codes of the patients from Aalborg.

10.2.2 Procedure

Before starting with the actual paradigm, the presence of a tactile P300 was affirmed with a simple tactile oddball run. Here, only left and right knees were stimulated, while participants always had to concentrate on the right knee (target frequency 0.25).

All participants performed three calibration runs to train a linear classifier, which was then used for three copy runs identical to those described in **study II** (concentrating on predefined positions, with feedback. See section 9.2.3). Patients in Aalborg did not see the monitor at this stage and were instructed verbally in their native language by the experimenter, who also provided feedback of the classification.

Finally, three guided navigation tasks similar to **study I** and Herweg et al. (2016) were performed.

After that, patients in Aalborg were finished with the session. Patient PT-N from Nürnberg additionally received a short questionnaire to record subjective data over several sessions.

10.2.3 Guided Navigation

Free navigation would have made it difficult to objectively evaluate BCI performance of patients, since it is often not clear whether a command was intentional or not. Instead, we modified the navigation paradigm to use a fixed course and only execute correct commands. Erroneous commands were received, but not executed and thus had to be repeated. The required course was indicated by a clearly visible line on the floor (see

Figure 10.1). The backrest of the chair was removed for better visibility of the line. Checkpoints were removed entirely.

We created three new courses, which could be navigated in both directions. Overall, this provided six different scenarios for navigation, all of which required ten successful commands to complete. After ten selections, the run was terminated and the course progress was recorded as a measure of accuracy.

The guided navigation was a further adaptation from a complete rework of the virtual environment which was done in another thesis (Bachelor of Engineering) outside the scope of the present dissertation (Eidel, 2018).

Figure 10.1: Updated environment for guided navigation. A line on the floor in front of the user indicates the required course. Numbers on the top of the screen show the total number of commands received and the total number of correct commands.

10.2.4 Questionnaires (Nürnberg)

The NASA-TLX was filled in digitally once a month at the end of the BCI session. The TUEBS was filled in once in session four.

To implement the UCD (see section 4.2), we included six short questions about the user experience at the end of all BCI sessions. As in Holz et al. (2015a), we used rating scales for the items *satisfaction, enjoyment, frustration,* level of *exhaustion*, subjective level of *BCI control*, and *control change* during the session (detailed in section 4.2.2). A set of these questionnaire items was applied and recorded digitally using a Python 2.7 based graphical user interface. Since this interface was intended to facilitate independent use by patients (potentially via eye-tracking), the interface was kept simple, with large, easily accessible buttons. Because of this, no *analogue* scales, but Likert-like rating scales with discreet values (integers or levels of the ordinal scales) were used.

10.2.5 Physiological Measures

Physiological measures were extracted from calibration data as described in section 7.2, using a rejection threshold of 75 μV and a time window of 300-500 ms post-stimulus. Values from electrode positions Fz, Cz and Pz were used for (descriptive) analysis and plots.

10.2.6 Statistical Analysis

Our sample was relatively heterogeneous due to the inclusion of patients in different stages of ALS. Because of this diversity and the small sample size, physiological measures from Fz, Cz and Pz were analysed descriptively only, and were not grand-averaged for a group-level analysis. Instead, the BCI was evaluated using a single-subject analysis based on the machine learning approach, since we assumed that this would better account for interindividual differences between the patients than a classical analysis of physiological measures extracted from predefined positions (when encountering high subject-to-subject variability, Blankertz et al. (2006) advocate the use of machine learning).

As in **study II**, the number of sequences (8), and thus the time required for one selection, was constant (as was the number of possible commands). Hence, ITR values depended only on accuracy. Because of this redundancy, accuracies were included in statistical analysis, while ITRs are reported only descriptively to facilitate comparisons to other studies.

10.3 Results

10.3.1 Aalborg

Physiology

EEG data at Fz and Cz from all patients are visualized in Figures 10.2 and 10.3. As previously in **study I** and **II**, deflections at Pz were smaller descriptively (see supplementary Figure A.6 for Pz).

The EEG signal from Patient 2 was contaminated by a considerable amount of noise, but visual analysis suggested a late P300 deflection nevertheless. Data from Patient 5 showed little (Cz) to no observable (Fz) deflections in the P300 range. Still, since this patient achieved relatively high accuracies, we analysed his data more extensively, including all twelve electrodes, but found no substantially higher deflections elsewhere. There were some notable amplitude spikes at about 20 and 240 ms in both target and non-target epochs, which seemed to coincide with the vibrotactile stimulation.

Patient 7 showed a positive deflection in the P300 range in both target and non-target epochs, with little difference between the classes. Patient 11 had a particularly large amplitude peak of 9.65 μV at 600 ms (Fz).

Table 10.2 provides an overview of all extracted values (amplitudes and curve differences) from the time window of interest. We ran exploratory Spearman tests to determine whether amplitudes or curve differences correlated with the ALFR-R scores, but found no significant result (all $p > .1$).

Table 10.2: Aalborg: Overview of extracted values [μV] averaged from the P300 window.

	Amplitudes			Difference		
Patient	Fz	Cz	Pz	Fz	Cz	Pz
2	0.63	0.41	-0.15	1.78	1.66	1.86
5	0.18	1.24	1.11	0.25	1.13	0.65
6	2.74	2.96	1.78	3.15	2.71	1.9
7	3.28	4.2	3.47	0.78	1.18	2.04
8	3.46	4.83	3.14	4.77	5.57	3.51
11	1.68	3.15	1.7	5.41	3.19	1.93
12	2.83	2.29	0.55	3.12	2.05	0.48

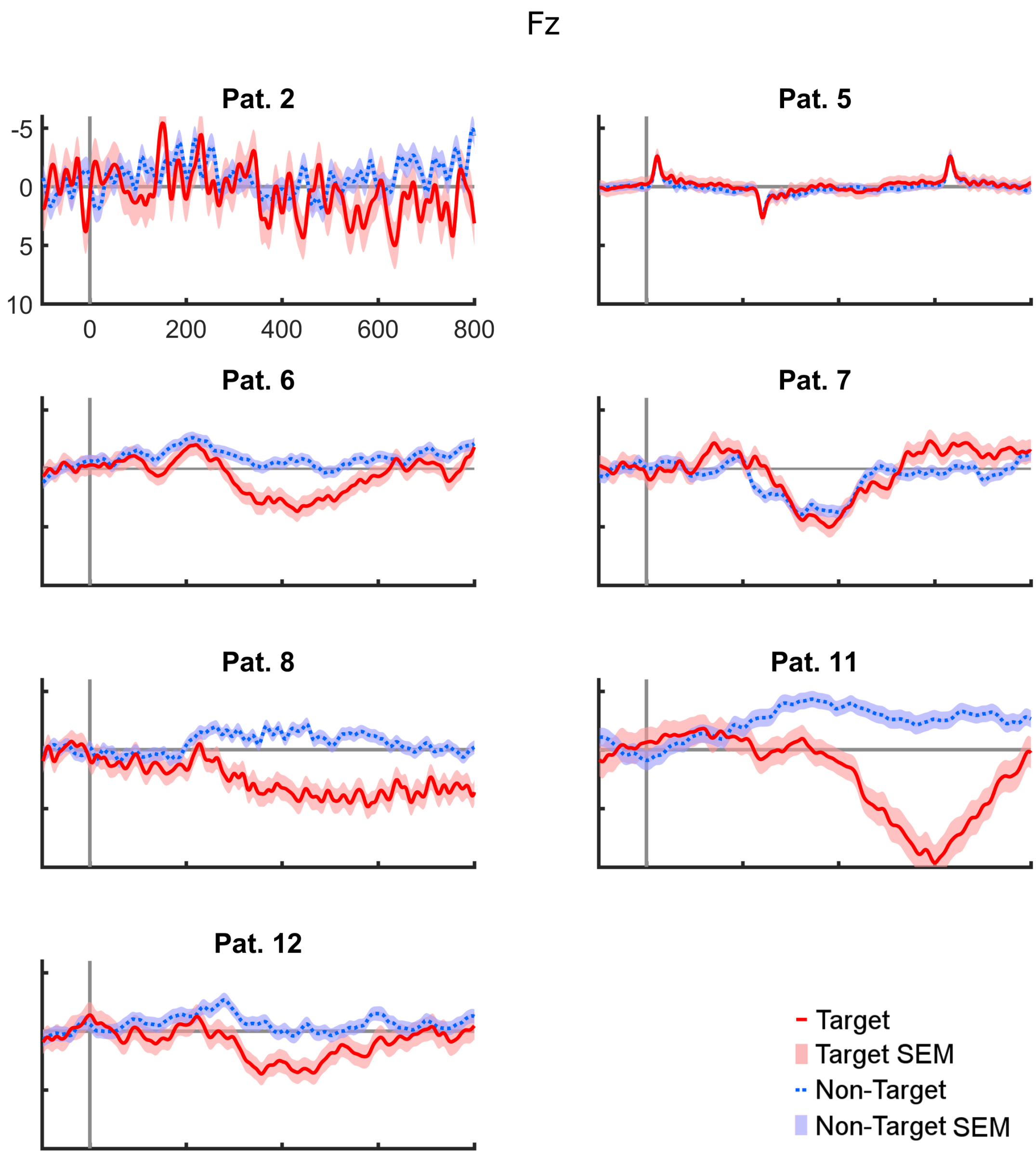

Figure 10.2: Post-stimulus target and non-target epochs at Fz. Data was averaged per individual. Pat. = Patient.

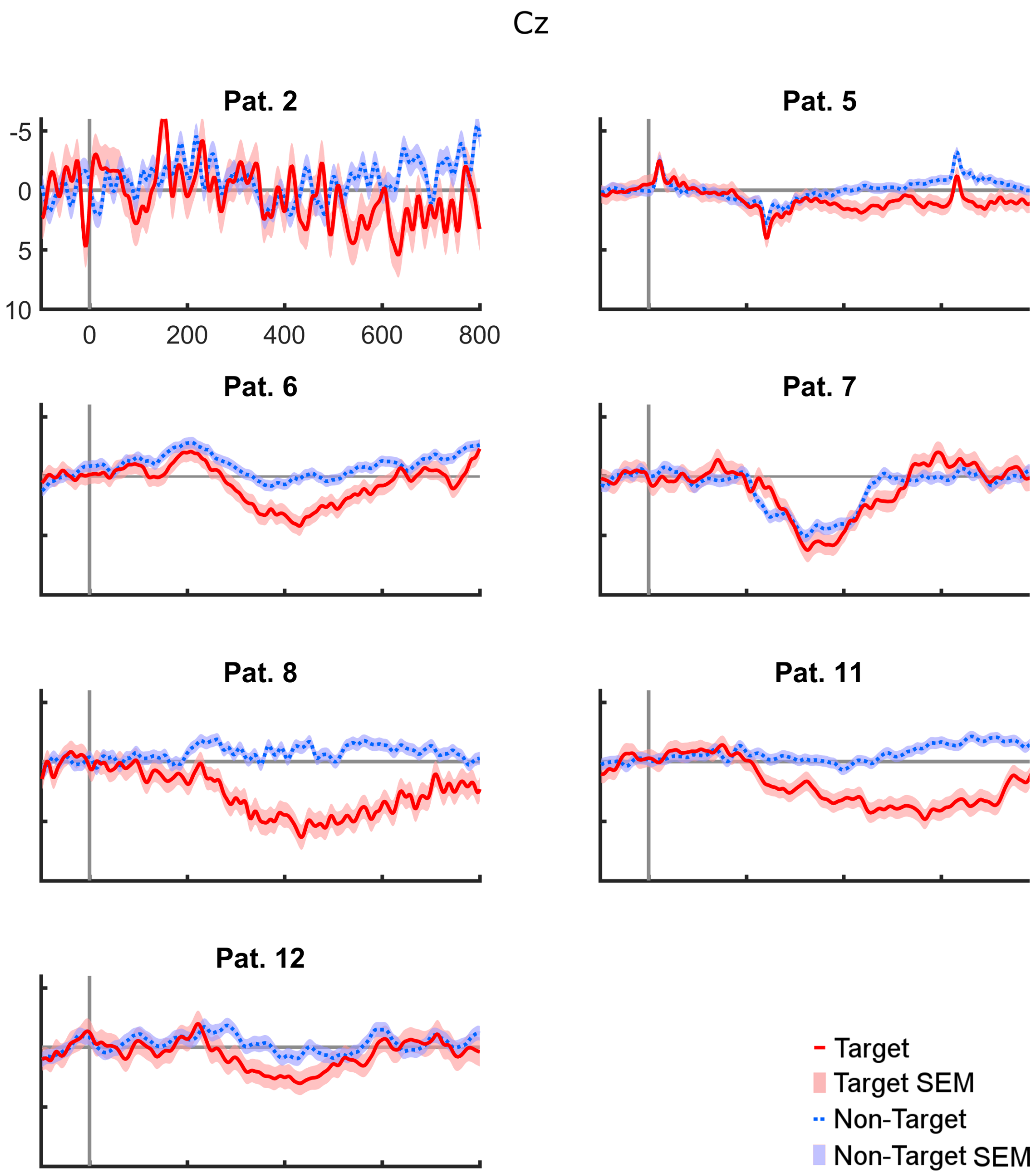

Figure 10.3: Post-stimulus target and non-target epochs at Cz. Data was averaged per individual. Pat. = Patient.

BCI Performance

Table 10.3 provides an overview of accuracies and ITR values from both copy and navigation tasks for all patients.

Accuracies from five patients were significantly or highly significantly above chance level in both tasks (approx. 40 % for $\alpha = .05$, or approx. 50 % for $\alpha = .01$) (Combrisson & Jerbi, 2015; Müller-Putz et al., 2008). Highest performances were achieved by Patients 2, 5, 8 and 11, who also performed above the criterion level of 70 % (Kübler et al., 2001b). On average, accuracies were higher in the copy task (57.7 %, SD = 21.2) than in the navigation task (39.8 % SD = 15.6) (t-test $p = .016$). Patient 7 achieved no control in either task. Again, we ran exploratory Spearman tests to determine whether accuracies correlated with the ALSFR-R scores, but found no significant results. Further, we found no correlation of amplitudes or differences between curves with copy or navigation accuracies (all $p > .1$).

Table 10.3: Aalborg: Overview of BCI performances (accuracy and ITR). Asterisks indicate significance over chance level (*<.05, **<.01)

	Copy		Navigation	
Patient	Accuracy [%]	bits/min	Accuracy [%]	bits/min
2	62.50**	1.36	50.00**	.63
5	83.33**	3.28	43.33*	.35
6	37.50	.17	40.00*	.24
7	25.00	0	15.00	0
8	70.83**	2.02	60.00**	1.19
11	75.00**	2.40	46.67*	.48
12	50.00**	.63	23.33	0

10.3.2 Nürnberg

By the end of 2019, the patient from Nürnberg (PT-N) had participated in 23 full sessions and filled in six complete NASA-TLX questionnaires. The TUEBS was filled in only once, as it was very straining and time consuming by the patient's own report.

Session one was excluded from analysis because the number of sequences was adjusted from six to eight to fit the patient's performance, and because of a misunderstanding of the navigation task. Physiological data from sessions two (all electrodes) and four (Fz only) were excluded due to a deviation of more than two standard deviations from the mean.

Physiology

For descriptive purposes, Figure 10.4 delineates the patient's global average ERPs, which includes all sessions (and thus does not account for changes over time).

Table 10.4 provides a summary of the respective extracted mean amplitudes and mean differences on Fz, Cz and Pz, revealing that overall, deflections from PT-N were highest at Fz and Cz.

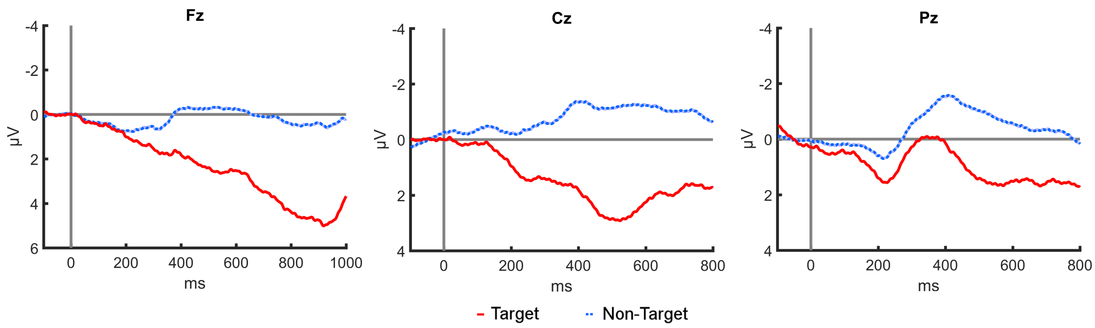

Figure 10.4: Global ERP shapes, averaged over all sessions. Note that for Fz, axis limits were extended to fit the ERP.

Table 10.4: Summary of mean amplitudes and mean differences, averaged over all sessions.

	Fz	Cz	Pz
Amplitude [μV] (SD)	2.19 (.97)	2.37 (.54)	.98 (.48)
Difference [μV] (SD)	2.77 (1.17)	3.45 (.67)	2.05 (.61)

Visual analysis of the EEG epochs across all sessions revealed that the patient was able to elicit the P300 reliably, exhibiting a clearly visible deflection in all sessions (see Figures 10.7 and 10.8 for Fz and Cz; see supplementary Figure A.7 for Pz). Data from session 10 showed a positive deflection despite being contaminated by a large amount of noise.

To visualize training effects on the P300, we plotted a pre/post comparison of EEG data. For this purpose, epochs from the first and last three included sessions were averaged (Figure 10.5). Descriptively, target amplitudes diminished slightly across the training sessions, while retaining the same approximate peak amplitude. Non-target amplitudes shifted toward the negative polarity over time.

Furthermore, we plotted the mean amplitudes of Fz, Cz and Pz across all included sessions and calculated a linear regression for each electrode (Figure 10.6). Descriptively, no notable effect of training could be observed. One-tailed Pearson correlations revealed no significant training effect for any electrode (performed for amplitudes and differences between curves; all $p > .1$).

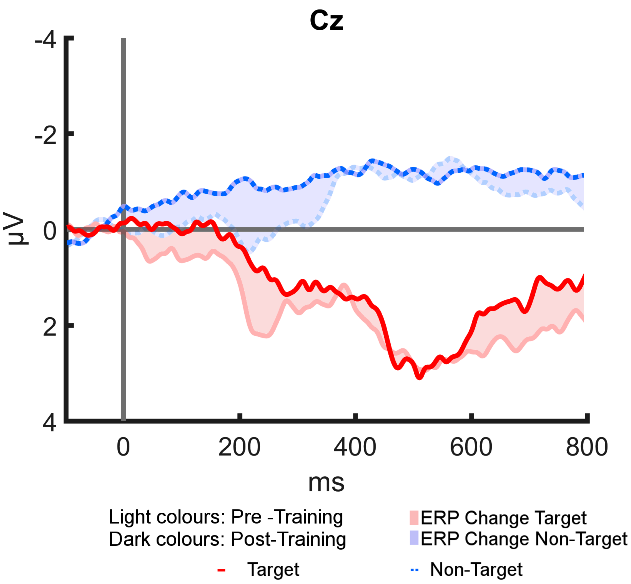

Figure 10.5: ERP changes between the averages of the three first and the three last included sessions.

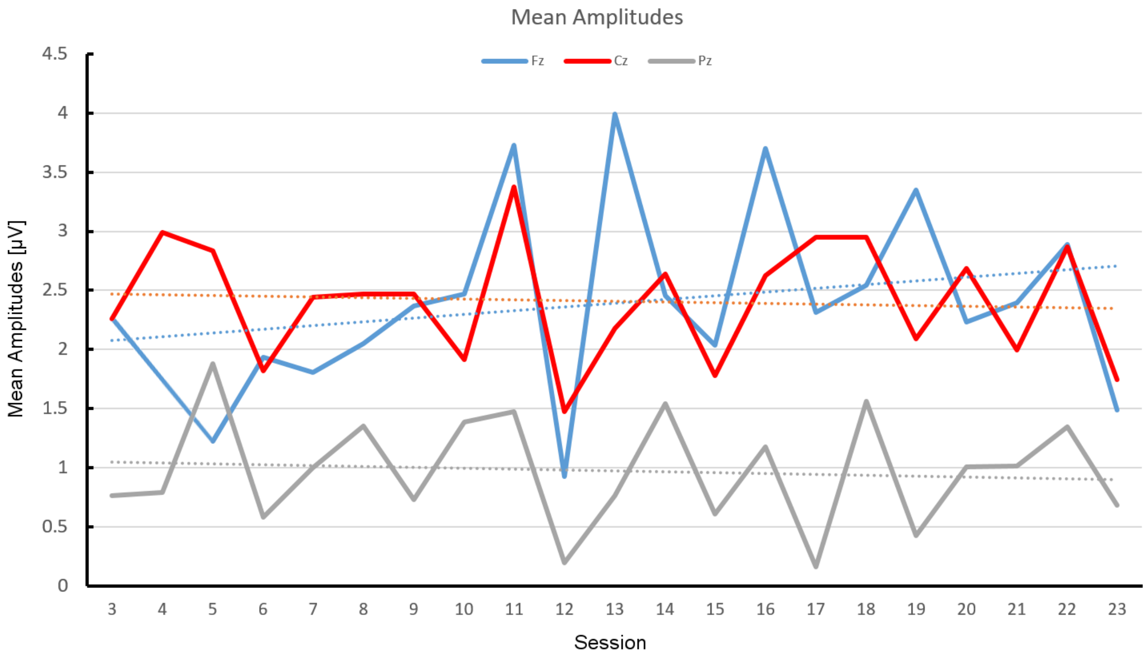

Figure 10.6: Mean amplitudes across all included sessions of PT-N. Dotted lines represent the linear regression.

BCI Performance

Accuracies in both copy and navigation tasks were always significantly above chance level (Combrisson & Jerbi, 2015; Müller-Putz et al., 2008), with the navigation task in session 23 being the only exception (Figure 10.9). Again, no significant Pearson correlation could be found between training session and accuracies, neither from the copy task (r_p = -.07, p = .621) nor the navigation task (r_p = -.10, p = .668).

Across all sessions, mean accuracies were 67.7 % (SD = 11.7) and 60.2 % (SD = 11.2) for copy and navigation tasks, respectively. An exploratory t-test revealed that this difference was significant (p = .006). Overall, the 70 % criterion was reached in the copy task in nine, and in the navigation task in four sessions.

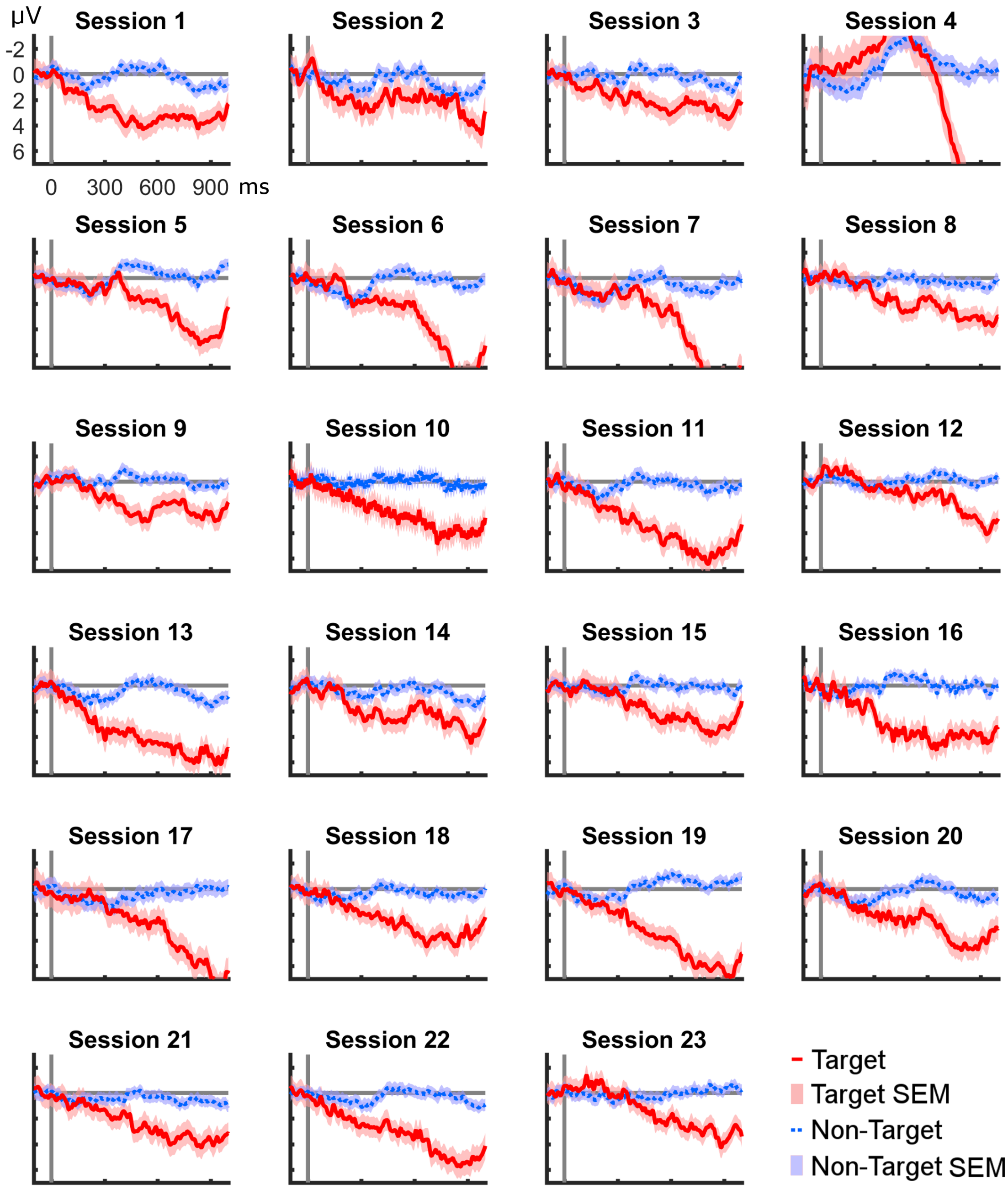

Figure 10.7: Post-Stimulus EEG (Fz) from PT-N, by session.

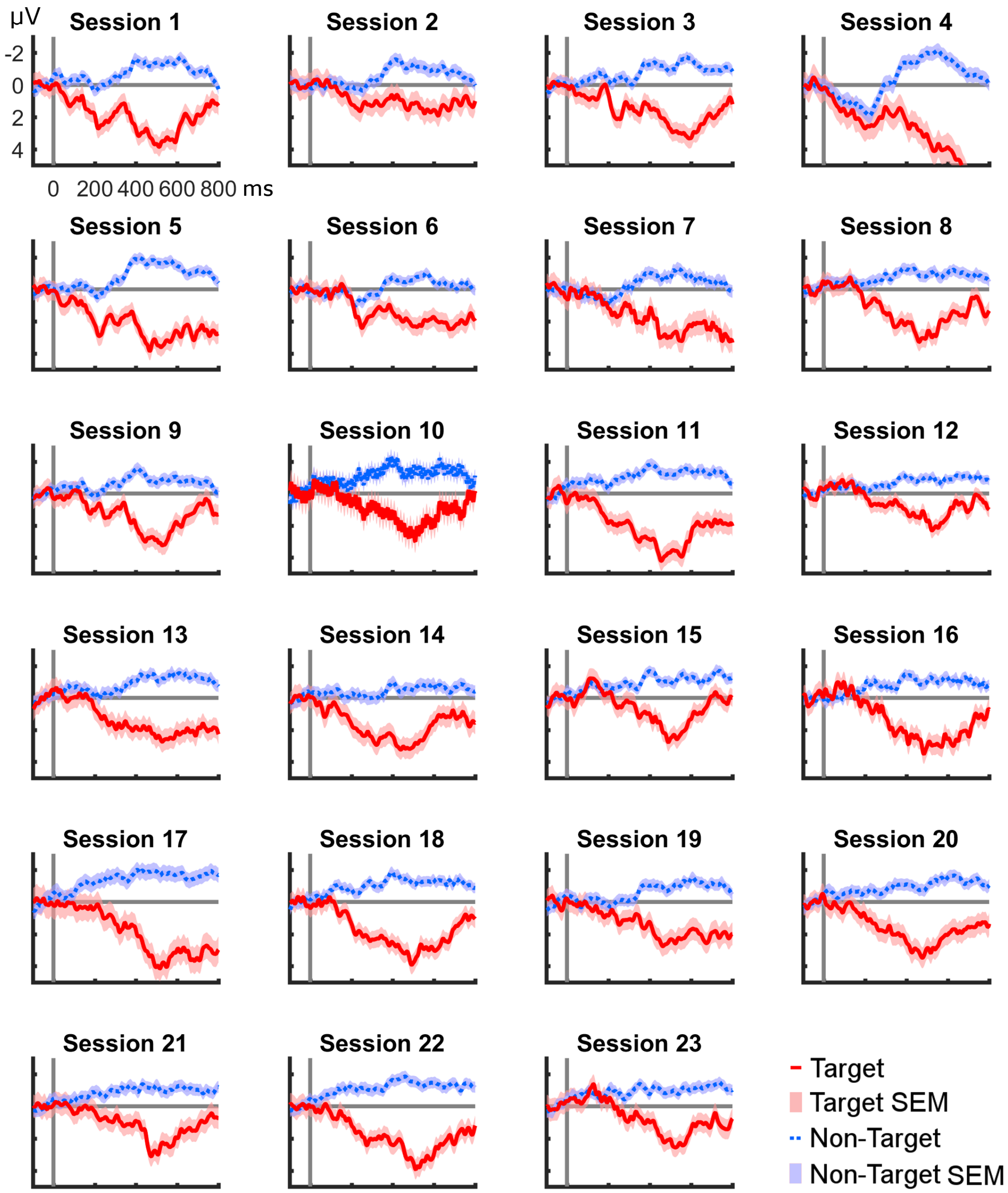

Figure 10.8: Post-Stimulus EEG (Cz) from PT-N, by session.

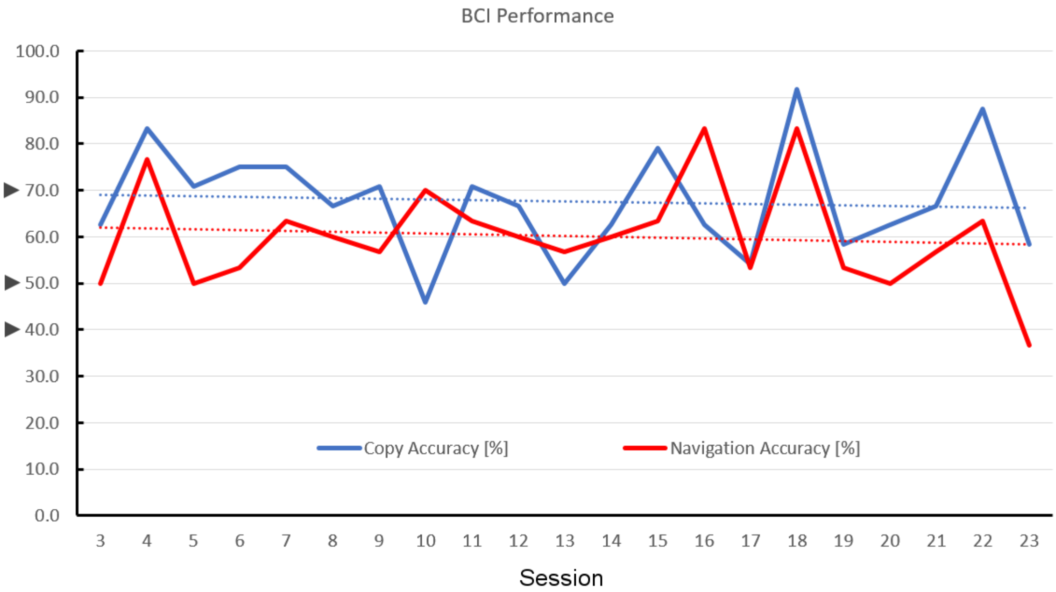

Figure 10.9: BCI performance from copy and navigation tasks across all included sessions of PT-N. Dotted lines represent a linear regression. Triangles indicate chance level thresholds of 40 % ($\alpha = .05$) and 50 % ($\alpha = .01$), and the 70 % criterion.

Rating Scales for User Experience

Figure 10.11 provides an overview of the ratings of the six questionnaire items that were applied after every session. On average, user satisfaction ($M = 4.86$, SD $= 2.15$), enjoyment ($M = 5.52$, SD $= 2.48$) and frustration ($M = 5.33$, SD $= 2.04$) were all rated at a medium level, but there was a relatively high variance between sessions. Exhaustion was rated medium in all but three sessions. Subjective control was rated medium to low. Regarding the change of control during a session, the patient reported no change in ten sessions, worse control in seven sessions and improved control in four sessions. Descriptively, no pattern over the time of the training program was observed for any of these measures. To test for hypotheses **H4a** and **H4b**, we calculated the correlations between satisfaction and accuracies and subjective control (see Figure 10.10). While we found no significant result for neither BCI accuracies from copy nor from navigation tasks (both $p > .1$), a highly significant correlation between satisfaction and subjective control was revealed ($r_p = .76$, $p < .001$). We found no correlation between subjective control and the objective measure of BCI accuracies (both $p > .1$).

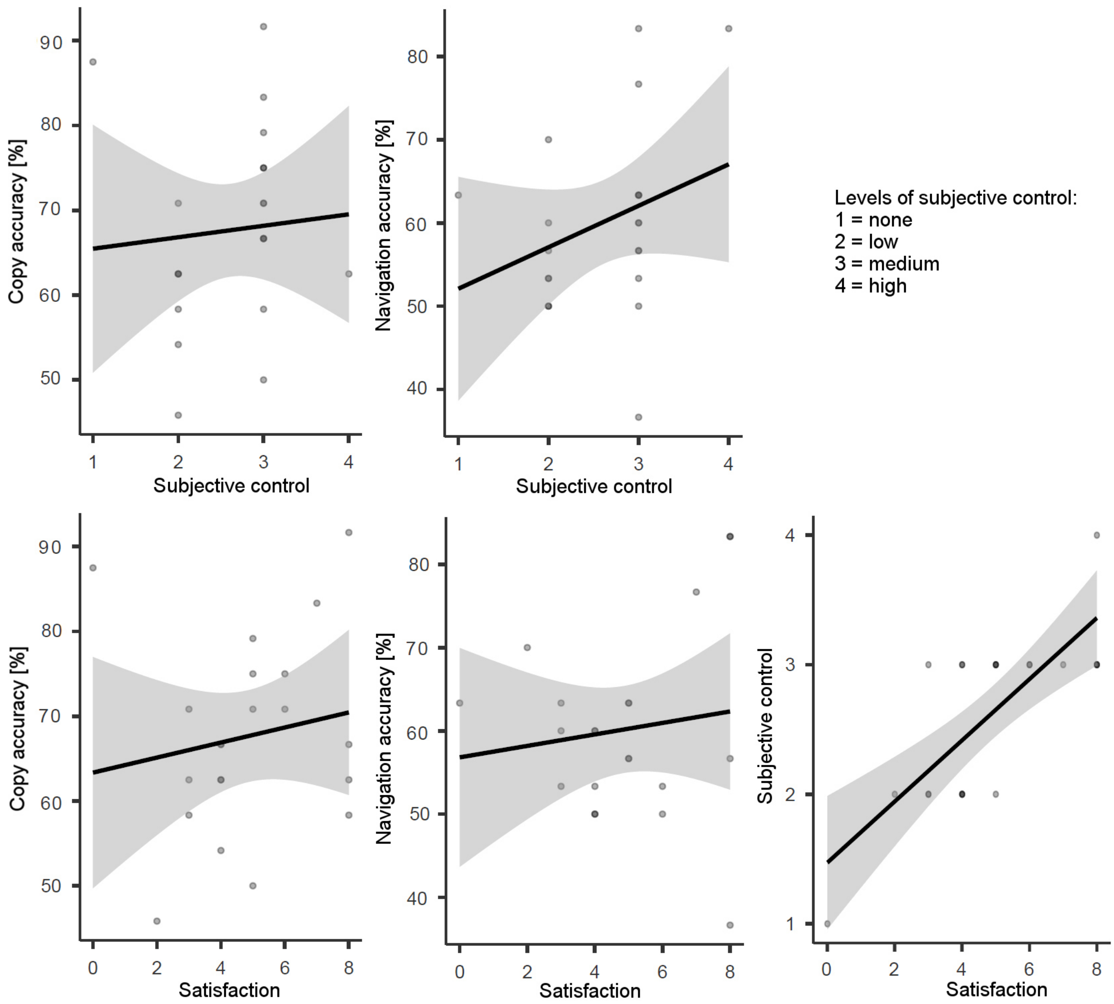

Figure 10.10: Correlation between subjective control / satisfaction and accuracies. Shaded area indicates the 95 % confidence interval.

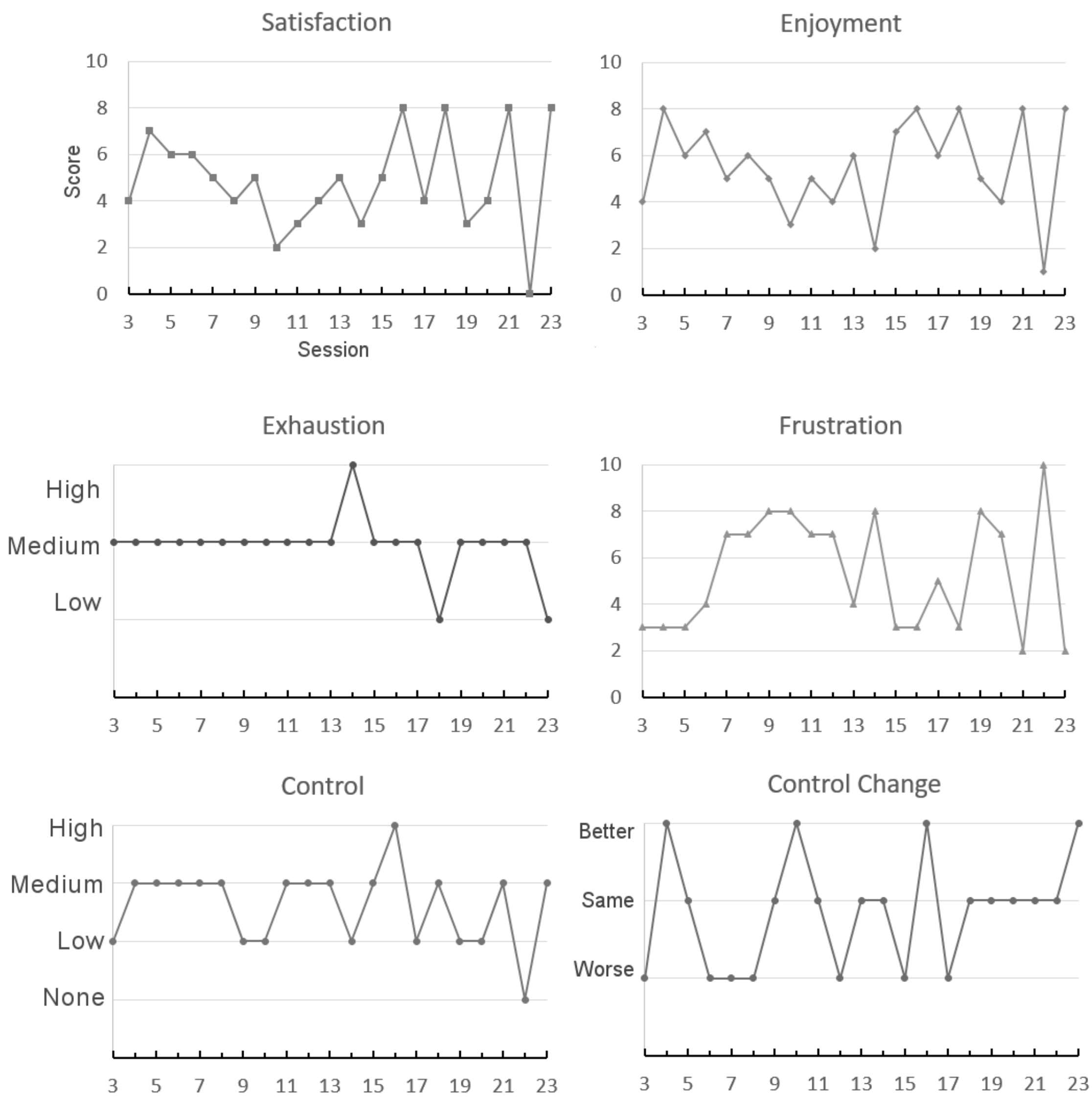

Figure 10.11: Subjective data from PT-N, recorded after every session.

NASA-TLX

Due to the small sample of only five iterations of the NASA-TLX, analysis was done descriptively. The NASA-TLX resulted in a mean total workload score of 44.7 (SD = 10.7), with the subscale *mental demands* (M = 12.2, SD = 5.0) being one of the highest contributors to the score, as compared to *physical demands*, which were scored very low (M = 2.0, SD = 2.1) on all five applications of the NASA-TLX (Figure 10.12). No significant training effects could be observed for the total score or *mental demands* (**H3a/b**). *Mental demands* appeared to peak in session eleven of the training program. Descriptively, *performance* appeared to increase, and *effort* to decrease over time. Frustration fluctuated a lot between sessions (M = 13.2, SD = 10.1). Table 10.5 provides a summary of all values from the NASA-TLX.

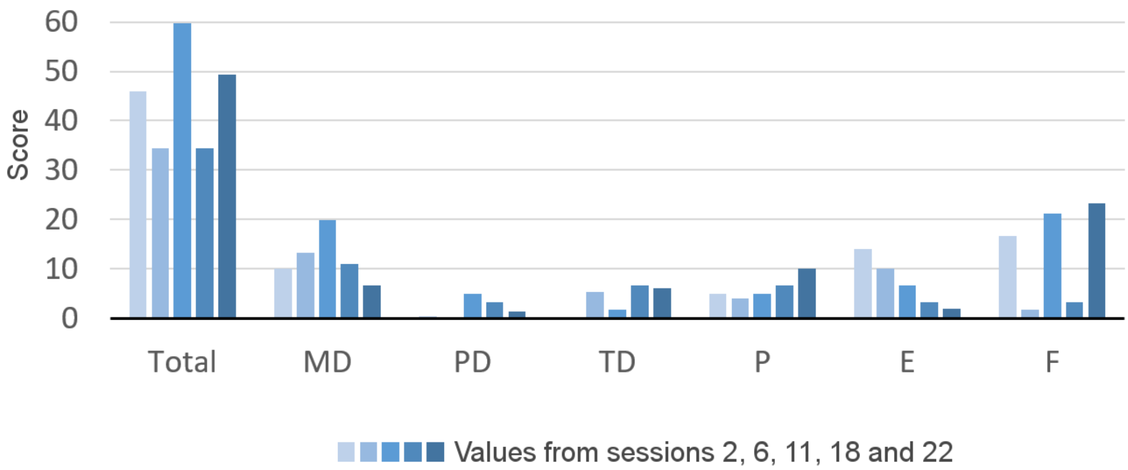

Figure 10.12: NASA-TLX data across sessions. MD = mental demands. PD = physical demands. TD = temporal demands. P = performance. E = effort. F = frustration.

TUEBS

PT-N completed the TUEBS once in session three. All ratings and comments are provided in Table 10.6.

The mean score for the main items was 2.25. The patient was not at all satisfied with the system's *ease of use* and *effectiveness*, and often remarked that its use was

Table 10.5: Nürnberg: NASA-TLX scores from five BCI sessions. MD = mental demands. PD = physical demands. TD = temporal demands. P = performance. E = effort. F = frustration. Total = global NASA-TLX score. Raw = unweighted global score.

Session	MD	PD	TD	P	E	F	Total	Raw
2	10.0	0.3	0.0	5.0	14.0	16.7	46.0	36.7
6	13.3	0.0	5.3	4.0	10.0	1.7	34.3	30.0
11	20.0	5.0	1.7	5.0	6.7	21.3	59.7	55.0
18	11.0	3.3	6.7	6.7	3.3	3.3	34.3	38.3
22	6.7	1.3	6.0	10.0	2.0	23.3	49.3	41.7
Mean	12.2	2.0	3.9	6.1	7.2	13.2	44.7	40.3
SD	5.0	2.1	2.9	2.4	4.9	10.1	10.7	9.2

not yet possible without a trained specialist. He was "more or less" satisfied with the *safety* and *comfort*. The score for the BCI items was 2.0. The patient was "more or less" satisfied with the system's *reliability*, but not at all satisfied with its *speed*. He often remarked that he did not yet feel experienced enough to adequately answer a question. In the last part of the questionnaire, the patient selected *comfort*, *reliability* and *speed* as the most important (out of twelve available items).

Table 10.6: PT-N: TUEBS ratings from session three. Comments by the patient were translated from German. Items 12a-12c were not rated.

Item	Rating	Comments
1. Dimensions	2	System is still a provisional solution
2. Weight	5	-
3. Ease of adjusting	2	Not usable without a specialist
3a. ...EEG prepping	4	Can be done by a trained person
3b. ...Hardware setup	1	Installation and setup requires specialists
3c. ...Software setup	1	Only possible for specialists
4. Safety	3	I can not judge this
5. Comfort	3	The cap is not really comfortable, but tolerable
6. Ease of use	1	Currently only usable with specialists
7. Effectiveness	1	System currently not fit for every day use
8. Service	1	Not relevant at the moment
Mean Score (Main Items)	**2.25**	
9. Reliability	3	No experience
9a. ...of the EEG	3	No experience
9b. ...of the Hardware	3	No experience
9c. ...of the Software	3	No experience
10. Speed	1	Speed is suboptimal due to the system
11. Learnability	2	I can only judge this with more experience
12. Aesthetic Design	-	System has currently no "ratable" design
Mean Score (BCI Items)	**2.0**	

10.4 Discussion

10.4.1 Aalborg

Physiology

There appeared to be a high variance between the patients. Still, visual analysis revealed the presence of at least a small P300 deflection in all participants. This is promising, as it suggests that despite ALS and its possible effects on the central nervous system and somatosensory perception (as outlined in sections 2.2 and 4.1), tactile stimulation is feasible, thus providing a basis for a tactile BCI modality for ALS users.

Amplitudes from patients 7, 8 and 11 exceeded the mean value from **study II** ($M = 3.13\ \mu$V) from young, healthy participants in their first session. Notably, with an ALSFR-R score of 25, patient 7 was the most impaired participant of this sample.

The prominent amplitude spikes found in the signal of Patient 5 are most likely electromagnetic stimulation artifacts caused by the tactors, since the two first spikes appear at an interval of approximately 220 ms, which was the stimulus duration. The third spike, therefore, was the start of the subsequent stimulation. Occurring outside of the P300 window, and in both target and non-target epochs, the artifacts should have had little effect on the classification performance.

BCI Performance

Patients 5, 8 and 11 exceeded the criterion level of 70 %, whereas the highest value in the navigation task was only 60.0 % (Patient 8). This might be due to an effect of exhaustion, since the task orders were not randomized (instead, we wanted to present a gradually increasing complexity of the task to the patients). Moreover, it could have been due to a higher difficulty of the navigation task. However, these putative negative effects might have been counterbalanced to a degree by within-session training effects, i.e. the patients having a higher proficiency and better understanding of the paradigm by the time of the last task.

Surprisingly, Patient 5 achieved the highest accuracy despite comparatively small P300 deflections. McCane et al. (2014) analysed a visual P300 BCI among a group of 25 ALS patients with severe impairments (ALSFRS-R Mean Score = 6). In this study, the authors found that P300 amplitudes were diminished, but reported that this did not necessarily affect BCI performance.

Patient 7 achieved no control over the system in either task, despite a large P300 amplitude in the target epochs. This is likely due to the fact that non-targets elicited a nearly identical deflection in the P300 range, making it impossible to discriminate the classes.

10.4.2 Nürnberg

Physiology (H1)

Encouragingly, the P300 could be clearly identified in every session at Fz, Cz and to a lesser degree at Pz. However, with mean amplitudes of 2.37 μV (Cz, averaged over all sessions), the patient's P300 amplitudes were considerably smaller than the mean amplitudes from healthy participants in **studies I** and **II**.

Against our expectation, we found no significant change of the P300 across 21 training sessions and thus no support for **H1**.

Performance (H2)

On average, performances in the copy task were significantly better than in the navigation task (67.7 % versus 60.2 %). As discussed above for the patients from Aalborg, this suggests either an effect of fatigue, or that the navigation was in fact, more difficult. While the patient was able to operate the BCI with accuracies above chance level, on average his performance did not reach the 70 % criterion. Since this did not improve with training, we could not confirm hypothesis **H2**.

BCI Training Effects: Summary

We found no evidence for training effects in physiology or performance measures. This is partially consistent with other studies of long term BCI use by LIS patients, most notably by Holz et al. (2015b) and Botrel et al. (2017), who found no consistent increases on amplitudes over long periods of training. Rather, they reported a progressive decrease (after an initial increase) of amplitudes for one patient, and no significant change for another patient. Moreover, they found that accuracies were not fully dependent on P300 amplitudes, as one patient's accuracy remained stable despite amplitude decreases, whereas the other patient's accuracy decreased despite stable amplitudes. Similarly, a study by Nijboer et al. (2008b) found no performance increase in ALS patients who used a P300 BCI over months. Healthy participants, on the other hand, are well documented to benefit from even a few training sessions (Baykara et al., 2016; Halder et al., 2016; Herweg et al., 2016).

A weekly session, as performed by PT-N, may have been too infrequent to cause observable training effects (Kübler et al., 2004). Hence, multiple sessions per week might be worth considering.

Further, **study II** provided evidence that training effects of the tactile BCI may be related to improvements along the somatosensory pathway, which, as discussed in section 4.1, can be affected in ALS at multiple locations - thus, an explanation for the lack of improvements among patients might be found here. However, the present data did not allow to (anatomically) locate this effect more precisely.

NASA-TLX: Workload (H3a) and Mental Demands (H3b)

The patient's rating of the workload averaged at a total score of 44.7, and thus bordered on the 40 % percentile which begins at 45.0 (Grier, 2015). Surprisingly, this was well below the average scores of 63.2 and 55.8 from **studies I** and **II**, respectively, but also below the score of other BCI studies with healthy (Käthner et al., 2013) or impaired participants (Riccio et al., 2015). The present low workload rating may be due a different subjective baseline of this LIS patient, who may experience a highly inflated workload with tasks that are trivial for the healthy population. Thus, the BCI's demands may have been judged relatively low in comparison to other challenges. Much like in our analyses of healthy participants however, *mental demands* were rated significantly higher than *physical demands*.

Despite analysing workload across many training sessions, we found no training-related changes of total workload or its *mental demands* subcomponent, and no manifestation that could be interpreted as an effect of automaticity (Logan, 1988). Our hypotheses **H3a** and **H3b** could therefore not be confirmed. This is consistent with our analyses of the P300 measures and BCI performance, in which no training effects were revealed.

Rating Scales for User Experience (H4a, H4b)

Several items were recorded after every session to explore the patient's subjective experience and, in the spirit of the UCD, to be able to react to possible emerging patterns (e.g. extreme frustration or lack of satisfaction). On average, none of the items received notably high or low ratings, and no clear pattern emerged.

We hypothesized that user satisfaction would correlate with accuracy (**H4a**), but found no supporting evidence. However, similar to the study by Holz et al. (2015b), a strong correlation between satisfaction and subjective control was revealed, confirming **H4b**. Since we found no correlation between subjective and objective control (i.e. accuracies), this measure was likely more affected by mood than BCI performance. Further, it became clear that subjective control was not redundant to objective accuracy measures, so it seems advisable to keep recording this metric in future studies.

TUEBS: Usability

The patient was not convinced of the applicability of the TUEBS to his specific case. Due to the high effort of completing this relatively long questionnaire, only one was available for analysis.

As the TUEBS was completed early during the training program, he reported that he did not yet feel experienced enough to adequately judge some items. Still, ratings revealed that the BCI system has to be improved on in several aspects, as overall scores were considerably lower than in comparable studies (e.g. Holz et al., 2015a; Zickler et al., 2011). This may be in part because for this first training phase, the patient received a paradigm for wheelchair control: According to a questionnaire about subjective experiences of ALS patients (Hecht et al., 2002), "loss of mobility" was reported to be among the worst aspects of the disease. However, during this study, the patient has expressed a wish for a paradigm for communication instead. Such a paradigm is currently in development and may be experienced as more satisfactory.

Most notably, the patient was unsatisfied with the system's *speed*, *ease of use*, *effectiveness* and *learnability*. On many items, he remarked that trained specialists were required to set up and operate the system. For an independent use of the BCI, family and caregivers have to be thoroughly instructed to operate the system. According to some definitions, this group is contained in the term "end-user" (Sullivan et al., 2018).

Consequently, to close the translational gap (Kübler et al., 2013), it seems advisable to include them in the iterative design process, while making operation as easy as possible.

The patient considered *comfort*, *reliability* and *speed* as the most important factors for satisfaction, which is mostly consistent with similar studies (Zickler et al., 2011). While in the near future, the EEG cap could be better adapted for this end-user, especially in regards to physical comfort, the system's overall performance should also receive more immediate attention for improvements.

10.5 Effects of ALS on P300

Our results were very heterogeneous between patients. While the P300 is generally considered an ERP which can be elicited reliably (Polich, 2007), it is also known that there are considerable interindividual differences in BCI performances (Allison & Neuper, 2010).

For both our single-case from Nürnberg as well as our sample from Aalborg, possible effects of ALS, specifically on P300 amplitudes, should be kept in mind. There is a substantial body of research that indicates that P300 amplitudes are decreased in ALS patients (Hanagasi et al., 2002; McCane et al., 2015; Paulus et al., 2002). In other cases however, no significant changes of amplitudes or latencies were found, for instance in a study by Erlbeck et al. (2017), or Silvoni et al. (2016) (who analysed the tactile P300). The present results are in apparent contradiction to the latter studies. However, with a mean ALSFR-R score of 28.3 (SD = 8.2) (Silvoni et al., 2016) and 21.8, (SD = 10.5) (Erlbeck et al., 2017), these participants were substantially less affected by ALS than PT-N (ALSFR-R = 0). Results, therefore, may not be fully comparable.

10.6 Conclusion

We provide both a longitudinal single-case, as well as a cross-sectional assessment of the same tactile BCI for virtual wheelchair control among ALS patients. The single-case study in particular is a novelty and gave valuable insights about the long-term usability of the tactile modality for potential end-users.

We could demonstrate that BCI operation above chance level was possible for most participating ALS patients without prior experience with BCI in at least one of the tasks. Thus, the feasibility of the tactile paradigm among ALS patients is supported, albeit not for all patients. Those who could not achieve control in their first session, however, might be able to do so after some training. Performances were better in the copy task than during navigation, likely because of an increased difficulty of navigation.

Overall, the fact that P300 and BCI performance of PT-N, a patient with a severe and *progressive* neurodegenerative disease, did not substantially decrease over time should be seen as encouraging. Nevertheless, this study design should be further optimized to facilitate potential training and improve overall performance. One approach might be to improve the tactors to provide stimulation with a higher intensity (provided this does not increase the difficulty of ignoring non-targets). Further, increasing the training frequency to two or even three sessions per week might be beneficial (Kübler et al., 2004).

Part III

General Discussion

11

Overview

The present thesis provides a comprehensive and hierarchical approach to answer several key questions about the P300 BCI paradigm of interest and about the tactile modality for BCI in general. Notably, data was obtained not only from abstract BCI tasks, but from a 3D simulation of wheelchair navigation, which could be considered as a realistic and plausible use-case.

Starting out from the promising observations of Kaufmann et al. (2014) and Herweg et al. (2016), we first endeavoured to replicate these results in **study I**, to solidify the evidence for training effects and the paradigm's general feasibility.

Study II then continued on this with a substantially increased sample size, while contributing first clues about the physiological origins of the observed training effects, which is an important first step to further optimize the paradigm in the future.

In parallel, **study III** provided important insights about the tactile BCI's feasibility among potential end-users with ALS. At the time of writing, the long-term study with the patient from Nürnberg is still ongoing, and may provide further crucial information about the effects of prolonged training, while also allowing to observe potential effects of disease progression.

11.1 Research Goals

At the beginning of this thesis (section 6, page 47), three major research goals were defined, which shall now be revisited. Table 11.1 provides a condensed overview of the principal contributions of each study.

The goals are briefly restated for better readability:

- **Goal 1: Replicate training effects and high performance**

Both **studies I** and **II** replicated the essential results from the original work of Herweg et al. (2016), although previous high performances were not quite reached.

Study I was unexpectedly affected by a high number of BCI inefficiencies, but training effects on BCI performance could be demonstrated among those who could operate the system.

Study II included substantially more BCI-efficient participants and provided significant evidence for training effects on the P300, BCI performance and mental workload, but not workload in general.

Overall, we have demonstrated that the present paradigm is among the best-performing approaches - including both tactile and auditory modalities. The average post-training efficiency (9.50 bits/min) of **study I** was the second highest of comparable paradigms, and only outmatched by Herweg et al. (2016) (see Table 5.2).

Study III did not reveal training effects in the single-case analysis of a LIS patient, but P300 and performance measures remained stable over several months. Considering the progressive nature of ALS, this consistency of performance over an extended time period is encouraging on its own (see Nijboer et al., 2008b, for similar results with a visual BCI), and an important factor for the continued usability of the BCI among end-users.

To summarize, **goal 1** was achieved for healthy participants. The data from ALS patients are highly valuable even if no training effects could be shown and may be used to optimize the design of this ongoing study. Although average performance levels from **studies I** and **II** were higher overall than those of ALS patients, it should be noted that there were several healthy participants who did not reach the relatively high level of BCI control of PT-N.

- **Goal 2: Identify factors for training effects**

Importantly, **study II** revealed that across five sessions, participants' somatosensory sensitivity had increased significantly. Guided by similar, but BCI-unrelated studies, we designed and implemented a new approach to assess the vibrotactile frequency discrimination threshold. Judging by the strong and unambiguous result, this design proved to be suitable and robust against certain uncontrolled confounding variables (such as clothing). With this experiment, we were able to identify somatosensory sensitivity as highly probable factor for training-related improvements.

While we demonstrated training effects on both performance and sensitivity, a causal relation was not conclusively proven. Still, the demonstration of a strong training effect on somatosensory sensitivity was an important step toward research **goal 2**, and should be confirmed in future studies.

Conversely, a training effect on the robustness against distraction, which might have been interpreted as an effect of automaticity, could not be demonstrated. This was partly due to the inconsistent nature of the results - **Study II** indicated a decrease of mental workload, but **studies I** and **III** did not. More specifically, **study II** implemented an experiment (*dual task*) which was designed to assess the effect of workload increase on BCI performance. We could show that, generally, performance of a second task negatively affected accuracies. However, the data did not indicate the training-induced decrease of this distractive effect, contradicting our expectation. Thus, evidence for a training-related effect on mental workload was, overall, not sufficient.

Nonetheless, automaticity does not have to be discarded as a possible factor for performance increase yet. Inconsistent results about mental workload and automaticity should be clarified with further experiments.

- **Goal 3: Establish usability (and training effects) among ALS patients**

Study III demonstrated that almost all ALS patients, most notably our long-term participant with LIS, could operate the tactile BCI with accuracies above chance or even above the 70 % criterion level (Kübler et al., 2001b; Müller-Putz et al., 2008).

Training effects could not be demonstrated - however, the results from PT-N revealed that his reasonably high level of BCI performance remained stable over several months of weekly training, and no notable habituation effects or persistent drops in performance were observed. Therefore, the paradigm could already be regarded as established for at least one end-user. Since this is a continued effort, there is a lot that still remains to be done, for instance, to improve the overall performance of the system, to increase the ease of use of the entire system toward a more autonomous use, and to transform the paradigm further into the practical everyday applications requested by the patient.

The cross-sectional part of **study III** suggests that these results may be transferable to the general population of ALS patients (and possibly people with similar impairments): Most participants were able to achieve above-chance control without any previous experience. Those who did not, on the other hand, might benefit from training (as established among healthy participants). To conclude, general feasibility and usability of the tactile modality among (potential) end-users was demonstrated, so that **goal 3** was achieved to a large degree.

Table 11.1: Principal results and their contributions to the research goals.

	Training Effects			Training Factors		Goal		
Study	*P300*	*Performance*	*Workload*	*Sensory*	*Automaticity*	1	2	3
I	-	+	-			±		
II	+	+	±	+	-	+	+	
III	-	-	-			-		+

Notes: Plus signs indicate results consistent with hypotheses or that contributed to a research goal. Minus signs indicate that no such effect or contribution was found. Plusminus signs indicate inconclusive observations. Cells are left blank if the item was not applicable.

- **Additional merits**

All studies implemented elements of the UCD and revealed important insights from the users' perspectives. In **study III**, the long-term participation of a LIS patient included the frequent collection of subjective data on the user's experience, a regular evaluation of perceived workload and one iteration of the TUEBS to evaluate usability. The collection of these data was a first, but crucial step toward future user-centered optimizations of the paradigm and toward a narrowing of the translational gap: Thanks to the inclusion of a (potential) end-user, future iterations of the paradigm may now be designed based on these highly relevant findings.

A tangential, but important merit of the presented research effort is that a novel tactor system was developed in cooperation with the TH Nürnberg (Schuster, 2018). This device is affordable (which has already allowed us to reproduce and distribute it to patients and clinics) and relatively easy to use - both of these factors are vital to our endeavour to establish the paradigm in the patients' own homes.

11.2 Limitations

The current paradigm allows for only four different commands, while most (visual) matrix-based spellers (Farwell & Donchin, 1988) provide 25 or 36. Stimulation in the vibrotactile modality is relatively slow: BCI2000's default visual flash time is 31.25 ms, whereas in the tactile paradigm, one stimulation took 220 ms, substantially increasing overall stimulation time. Both of these issues severely limit the maximum possible ITR. However, both the C2 tactors and the Arduino-based prototype by Schuster (2018) can easily be extended to more than four classes or body positions - but the congruence of the command and the body position (Thurlings et al., 2012b) could be lost in the process.

Tactile stimulation is difficult to control for equal intensity, as there are several factors that influence its perception (tightness of attachment, clothing). To equalize stimulus intensity, using the participant's subjective report of stimulus intensity is inevitable, but this may have introduced a bias, for instance when a tactor was felt much stronger than the rest despite best efforts to the contrary. Another confounding factor may have been introduced by small differences in the neural transmission times of proximal versus distal stimulation positions, such that a signal from the knee arrived slightly later as compared to other positions. Sounds from the tactors may have contributed to ERP elicitation, which was not controlled for.

Furthermore, participation in **studies I** and **II** included five sessions and took considerable effort, sometimes without the possibility of monetary compensation. The total number of participants was thus relatively low, leading to low statistical power and, likely, random sampling effects. Ideally, more participants would have attended more training sessions, but this was not yet realized due to the aforementioned practical limitations. The cross-sectional part of **study III**, too, was affected by a small sample size.

Data from healthy participants are valuable for basic research, but it is not clear to which extent their results are transferable to ALS patients. Findings on the effect of ALS on the (tactile) P300 are still inconclusive (as discussed in section 10.5), but we do know that the somatosensory pathways can be negatively affected (section 4.1).

Moreover, our single case LIS training study, though most valuable in itself, may not be representative of the entire target group. It should also be noted that potential end-users include all LIS patients, not only those with ALS, although the present thesis was focused on the latter. In fact, BCI efficiency varies even among the healthy (Allison & Neuper, 2010), and this effect may be further amplified in the heterogeneous group of (ALS) patients. This could be partially alleviated by a cross-sectional study as in **study III**, provided the sample size can be increased.

The single case study was performed at the patient's own home, where confounding factors, such as warning noises from medical equipment or any other sounds from adjacent rooms, may have influenced the results. Although **study II** revealed that many healthy participants could operate the tactile BCI despite distraction, there is plenty of evidence that a strain on working memory can negatively affect the P300 amplitudes (Kok, 2001; Polich & Kok, 1995).

For ALS patients, the order of copy and navigation tasks was not randomized. This was done to gradually increase task complexity, but may have introduced confounding order effects due to fatigue. Although sessions were not overly long or effortful (according to the NASA-TLX ratings by PT-N), the attention span of some ALS patients may be limited to a relatively short time window (Hanagasi et al., 2002).

Finally, the included studies did not record the participant's motivation levels, nor implement an experimental modulation of motivation. Thus, no effects of motivation could be analysed, although some studies had reported such (Baykara et al., 2016; Kleih et al., 2010). Conversely, effects of motivation are still inconclusive, as they could not always be observed (Käthner et al., 2013; Simon et al., 2015).

12

Conclusion

This thesis provides an important contribution to the research on tactile P300-based BCI paradigms, and gaze-independent approaches in general. To date, the tactile modality remains rarely used for BCI, in part because of its lack of speed, as compared to more established visual paradigms. Only few other studies have attempted to use a tactile P300-based modality with ALS patients, and to our knowledge, we provide a unique and entirely novel report of the long-term home use of a tactile BCI by a LIS patient.

Continuing on the work of Kaufmann et al. (2014) and Herweg et al. (2016), the three studies presented in this thesis have repeatedly confirmed the paradigm's feasibility for operating certain applications, using the example of virtual wheelchair navigation. Moreover, the demonstrated level of BCI control could be considered as effective and highly efficient as compared to other non-visual approaches - tactile paradigms in particular (Table 5.2).

Locomotion, in fact, may be an application for which EEG-based BCI is exceptionally well suited due to the technology's high transportability. Indeed, first steps to provide a proof-of-concept for a BCI-controlled, physical wheelchair in real life have already been undertaken (Halder et al., 2017). However, the actual use of such mobile platforms by impaired patients via BCI is still far away, and arguably faces plenty of engineering and possibly legal challenges that are outside of the scope of this thesis. For instance, the wheelchair would have to implement sophisticated algorithms to avoid accidents. Even then, legal restrictions must be thoroughly examined to determine whether a paralysed user can be cleared for autonomous control of the device in public spaces. Home use, one the other hand, should be quite possible.

With the three studies presented in this manuscript, we have demonstrated that non-visual modalities, such as auditory or tactile paradigms, can be an important alternative to the otherwise more efficient visual paradigms, which may fail entirely if gaze or vision are impaired. Auditory paradigms have already been shown to be feasible (e.g. Käthner et al., 2013), but occupy the important auditory channel, requiring almost undivided attention during stimulation. The tactile paradigm should now be considered as another, valuable option for gaze-dependent BCIs: Since the visual and auditory channels remain free, the users experience no additional restrictions on their daily routines, media consumption or social interactions.

It should be emphasized that of several different BCI approaches which were offered to PT-N, he preferred the tactile paradigm as the best option for his future long-term home use. A single-case study by Kaufmann et al. (2013), who compared the use of visual, auditory and tactile paradigms with another LIS patient came to a similar conclusion and described the tactile modality as clearly superior.

Most importantly however, tactile (and auditory) paradigms could be some of the last possible resources that allow communication as patients enter the CLIS and become unable to use visual BCIs or any other AT. Thus far, no convincing means of communication with CLIS patients has been established with any BCI approach (Kübler, 2019). Even the visual modality, despite years of research and its superior performance (cf. non-visual modalities), was never successfully used by a CLIS patient. These potential end-users, however, may be the most important of all.

If BCI proficiency is learned early on and continuously trained during the progression of ALS, control might be retained into the state of CLIS (Mayaud et al., 2016). Encouragingly, training effects were demonstrated for healthy participants in **studies I** and **II**. In the single-case study of a LIS user (**study III**), physiological and performance measures did not increase, but could be held stable over an extended time period, which is an important prerequisite for long-term independent home use.

In conclusion, our results indicated that the presented tactile approach is feasible for the regular use by LIS patients - thus, we strongly encourage further research and the early involvement of potential end-users.

12.1 Implications and Recommendations

- The tactile modality appears to be relatively difficult, possibly because in daily life, the somatosensory channel is less important and well-trained than the visual. It could be shown, however, that it is trainable, illustrating the importance of multi-session designs and the early inclusion of patients to reach the tactile BCIs full potential.

- Despite a high baseline difficulty, the tactile paradigm turned out to be more robust against distraction than estimated during pre-experiments, and consequently, some participants were not negatively affected in their dual task performance. Thus, BCI operation may still work to a satisfactory degree even while - as demonstrated - attentively listening to a story or conversation, or while performing other, sometimes inevitable, multitasking activities of daily life. This could be another important factor for successful, independent home use.

- Evidence from other studies suggests that the congruence of the stimulus positions and the encoded direction was an important factor for the high effectiveness of our paradigm (Thurlings et al., 2012b). Unlike in many other non-visual paradigms (Simon et al., 2015; Waal et al., 2012), with our paradigm, no visual aid was required and no code had to be learned. Participants were able to understand the relationship between the stimuli and the commands easily. Because of these reasons, we consider the paradigm to be very intuitive. If the number of tactor positions will be increased in the future, a similar congruence should be retained, as far as possible.

- Other than ensuring congruence, there is little room to make the paradigm easier, as it consists of only four commands which require basically no memorization. The paradigm can be further reduced for simple binary (e.g. yes/no) communication, as previous studies have already explored (e.g. Ortner et al., 2014; Ziebell et al., 2019).

- Somatosensory perception may be important for training the tactile paradigm, but may be impaired among potential end-users. Thus, stimulus positions should be individually chosen to ensure that the user can perceive them as best as possible. In the same vein, stimulus intensities should be reasonably high, and developing new hardware that offers stronger stimuli may be beneficial. On the other hand, we suspect that difficulties in target/non-target discrimination may complicate classification, and this issue might be aggravated with stronger stimulation. Either way, it seems likely that improvements of sensory sensitivity at least partially account for the observed training effects.

- To narrow the translational gap, patients as well as caregivers should be included in the design process of the paradigm. Ideally, the role of the caregiver should be minimized to starting the application. Therefore, the entire system should be made easily accessible for end-users:

  The electrode setup could be reduced to the minimum of positions required for effective control (depending on the individual user), and an EEG system that is easy to prepare seems advisable. Moreover, patients and caregivers should not have to deal with the configuration of the BCI parameters. Rather, a simple, graphical user interface should be provided that offers the basic functionalities such as choosing, starting and stopping a (pre-optimized) paradigm. Brain Painting is an excellent example of this concept, as it implements a user interface which is easy to operate and that offers an automatization of important steps (e.g. calibration) of the BCI cycle (Botrel et al., 2015).

12.2 Future studies

While the general possibility of training the tactile paradigm could now be seen as adequately established, the extent of trainability has not been fully explored. If possible, future studies on healthy participants should include a longer training program - thus far, no clear consolidation of amplitude increases was observed in the five-session designs of the tactile paradigm.

Furthermore, it is important to replicate our results regarding somatosensory sensitivity increases and effects of distraction on performance. The role of the somatosensory system for tactile BCI training is an entirely novel finding that should be further investigated. For instance, because of the prevailing issue of BCI inefficiency, many studies have attempted to identify predictors for BCI performance in the recent years (e.g. Halder et al., 2013; Hammer et al., 2014; Kleih & Kübler, 2015). Although a correlation between sensory thresholds and BCI performance could not be demonstrated in **study II**, it may be worth to collect more data to determine whether the individual sensory thresholds have any predictive function on BCI performances.

Moreover, the paradigm was considered as relatively difficult, despite an arguably simple design that includes only four possible choices. If a further simplification of the paradigm turns out to be unfeasible, other means to enhance performance could be explored. Motivation, for instance, might be improved by the implementation of aspects of *gamification* - i.e. the integration of gaming elements and designs in a context unrelated to games. This does not imply using the BCI recreationally (although Brain Painting has been popular among end-users). Rather, practical applications, which were usually preferred by our patients, could be designed to be more engaging and entertaining with this concept (Škola et al., 2019). Many other recommendations (and known issues) of BCI paradigm design can be reviewed in the publications by Lotte et al. (2013) or Chavarriaga et al. (2017).

A large part of future research efforts should directly include potential end-users. As our research on ALS patients continues, in particular in cooperation with our ALS patient in Nürnberg (PT-N), we must endeavour to swiftly and dynamically react to the user's - possibly changing - individual requirements, as called for by the UCD. This is not only a scientific, but also an engineering challenge.

End-user PT-N specifically requested a tactile paradigm for basic communication. For a first prototype (designed in collaboration with the patient), tactor numbers are planned to be increased from four to six. To further increase the number of possible commands, a two-layer design is currently under development. Here, the first layer would provide a selection of six categories of related commands. After selecting a category, a second layer would offer five specific statements or commands, along with an undo option that returns to the first layer.

While the details of the next application are not yet set in stone, PT-N does wish to continue his participation. Should his state progress further toward the CLIS, hopefully the training program will have helped to achieve the major goal we have set together at the beginning: A prolonged ability to communicate with the help of the tactile BCI.

References

Ahn, S., Kim, K. & Jun, S.C. (2016). Steady-State Somatosensory Evoked Potential for Brain-Computer Interface—Present and Future. *Frontiers in Human Neuroscience*, **9**.

Allison, B.Z. & Neuper, C. (2010). Could Anyone Use a BCI? In D.S. Tan & A. Nijholt, eds., *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction*, Human-Computer Interaction Series, 35–54, Springer London, London.

Anderson, J.R. (2000). *Learning and memory: An integrated approach, 2nd ed.* John Wiley & Sons Inc, Hoboken, NJ, US.

Andrews, K., Murphy, L., Munday, R. & Littlewood, C. (1996). Misdiagnosis of the vegetative state: retrospective study in a rehabilitation unit. *BMJ*, **313**, 13–16.

Bauer, G., Gerstenbrand, F. & Rumpl, E. (1979). Varieties of the locked-in syndrome. *Journal of Neurology*, **221**, 77–91.

Baykara, E., Ruf, C.A., Fioravanti, C., Käthner, I., Simon, N., Kleih, S.C., Kübler, A. & Halder, S. (2016). Effects of training and motivation on auditory P300 brain–computer interface performance. *Clinical Neurophysiology*, **127**, 379–387.

Bede, P., Elamin, M., Byrne, S., McLaughlin, R.L., Kenna, K., Vajda, A., Pender, N., Bradley, D.G. & Hardiman, O. (2013). Basal ganglia involvement in amyotrophic lateral sclerosis. *Neurology*, **81**, 2107–2115.

Bell, J., Bolanowski, S. & Holmes, M.H. (1994). The structure and function of Pacinian corpuscles: a review. *Progress in neurobiology*, **42**, 79–128.

Berger, H. (1929). Über das elektrenkephalogramm des menschen. *European archives of psychiatry and clinical neuroscience*, **87**, 527–570.

Betts, J.G., Young, K.A., Wise, J.A., Johnson, E., Poe, B., Kruse, D.H., Korol, O., Johnson, J.E., Womble, M. & DeSaix, P. (2013). *Anatomy and Physiology*, vol. 2. OpenStax, Houston, Texas.

Birbaumer, N. & Cohen, L.G. (2007). Brain–computer interfaces: communication and restoration of movement in paralysis. *The Journal of Physiology*, **579**, 621–636.

Blankertz, B., Dornhege, G., Lemm, S., Krauledat, M., Curio, G. & Müller, K.R. (2006). The Berlin brain-computer interface: Machine learning based detection of user specific brain states. *J. UCS*, **12**, 581–607.

Bolanowski Jr, S.J. & Zwislocki, J.J. (1984). Intensity and frequency characteristics of Pacinian corpuscles. I. Action potentials. *Journal of neurophysiology*, **51**, 793–811.

Borasio, G.D., Pongratz, D.E. & (fuer die neuromuskulaeren Zentren der Deutschen Gesellschaft für Muskelkranke (DGM)) (1997). Gedanken zur Aufklärung bei amyotropher Lateralsklerose (ALS). *Der Nervenarzt*, **68**, 1004–1007.

Bortz, J. & Schuster, C. (2011). *Statistik für Human-und Sozialwissenschaftler: Limitierte Sonderausgabe*. Springer-Verlag.

Bosch, E.P., Yamada, T. & Kimura, J. (1985). Somatosensory evoked potentials in motor neuron disease. *Muscle & Nerve*, **8**, 556–562.

Botrel, L., Holz, E.M. & Kübler, A. (2015). Brain Painting V2: evaluation of P300-based brain-computer interface for creative expression by an end-user following the user-centered design. *Brain-Computer Interfaces*, **2**, 135–149.

Botrel, L., Holz, E.M. & Kübler, A. (2017). Using Brain Painting at Home for 5 Years: Stability of the P300 During Prolonged BCI Usage by Two End-Users with ALS. In D.D. Schmorrow & C.M. Fidopiastis, eds., *Augmented Cognition. Enhancing Cognition and Behavior in Complex Human Environments*, vol. 10285, 282–292, Springer International Publishing, Cham.

Brooks, B.R., Sanjak, M., Ringel, S., England, J., Brinkmann, J., Pestronk, A., Florence, J., Mitsumoto, H., Szirony, K. & Wittes, J. (1996). The amyotrophic lateral sclerosis functional rating scale-Assessment of activities of daily living in patients with amyotrophic lateral sclerosis. *Archives of neurology*, **53**, 141–147.

Brouwer, A.M. & Van Erp, J.B.F. (2010). A tactile P300 brain-computer interface. *Frontiers in Neuroscience*, **4**.

Brownell, B., Oppenheimer, D.R. & Hughes, J.T. (1970). The central nervous system in motor neurone disease. *Journal of Neurology, Neurosurgery & Psychiatry*, **33**, 338–357.

Brunner, P., Joshi, S., Briskin, S., Wolpaw, J.R., Bischof, H. & Schalk, G. (2010). Does the 'P300' speller depend on eye gaze? *Journal of Neural Engineering*, **7**, 056013.

Byers, J.C. (1989). Traditional and raw task load index (TLX) correlations: are paired comparisons necessary? *Advances in Industrial Erfonomics and Safety l: Taylor and Francis.*.

Cauna, N. (1964). The effect of aging on the receptor organs of the human dermis. *Advances in biology of skin*, **6**, 63–96.

Cedarbaum, J.M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., Nakanishi, A., Group, B.A.S. & Group, A.c.l.o.t.B.S. (1999). The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. *Journal of the neurological sciences*, **169**, 13–21.

Chatrian, G.E., Petersen, M.C. & Lazarte, J.A. (1959). The blocking of the rolandic wicket rhythm and some central changes related to movement. *Electroencephalography and clinical neurophysiology*, **11**, 497–510.

Chavarriaga, R., Fried-Oken, M., Kleih, S., Lotte, F. & Scherer, R. (2017). Heading for new shores! Overcoming pitfalls in BCI design. *Brain-Computer Interfaces*, **4**, 60–73.

Chen, M.L., Fu, D., Boger, J. & Jiang, N. (2019). Age-Related Changes in Vibro-Tactile EEG Response and Its Implications in BCI Applications: A Comparison Between Older and Younger Populations. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **27**, 603–610.

Cohen, J. & Polich, J. (1997). On the number of trials needed for P300. *International Journal of Psychophysiology*, **25**, 249–255.

Cohen-Adad, J., Mendili, M.M.E., Morizot-Koutlidis, R., Lehéricy, S., Meininger, V., Blancho, S., Rossignol, S., Benali, H. & Pradat, P.F. (2013). Involvement of spinal sensory pathway in ALS and specificity of cord atrophy to lower motor neuron degeneration. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, **14**, 30–38.

Combrisson, E. & Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of Neuroscience Methods*, **250**, 126–136.

Comerchero, M.D. & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, **110**, 24–30.

Daltrozzo, J. & Conway, C.M. (2014). Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Frontiers in human neuroscience*, **8**, 437.

Davis, H. (1961). Some principles of somatosensory action. *Physiol Rev*, **41**, 319–416.

Debaere, F., Wenderoth, N., Sunaert, S., Van Hecke, P. & Swinnen, S.P. (2004). Changes in brain activation during the acquisition of a new bimanual coordination task. *Neuropsychologia*, **42**, 855–867.

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

Demers, L., Weiss-Lambrou, R. & Ska, B. (2002). The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0): An overview and recent progress. *Technology and Disability*, **14**, 101–105.

Dharmadasa, T. & Kiernan, M.C. (2018). Riluzole, disease stage and survival in ALS. *The Lancet Neurology*, **17**, 385–386.

Donchin, E., Ritter, W. & McCallum, W.C. (1978). Cognitive psychophysiology: The endogenous components of the ERP. *Event-related brain potentials in man*, **349**, 411.

Duncan, C.C., Barry, R.J., Connolly, J.F., Fischer, C., Michie, P.T., Näätänen, R., Polich, J., Reinvang, I. & Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, **120**, 1883–1908.

Eidel, M. (2018). Entwicklung eines Tools für simulierte Rollstuhlsteuerung per Brain Computer Interface (BCI) (Unpublished Bachelor Thesis), Julius-Maximilians-Universität Würzburg.

Erlbeck, H., Mochty, U., Kübler, A. & Real, R.G. (2017). Circadian course of the P300 ERP in patients with amyotrophic lateral sclerosis-implications for brain-computer interfaces (BCI). *BMC neurology*, **17**, 3.

Erp, J.B.F.v. & Brouwer, A.M. (2014). Touch-based Brain Computer Interfaces: State of the art. In *2014 IEEE Haptics Symposium (HAPTICS)*, 397–401.

Faller, J. (2012). BCIs that use steady-state visual evoked potentials or slow cortical potentials. In *Brain-Computer Interfaces, Principles and Practice*, Oxford University Press.

Faller, J., Müller-Putz, G., Schmalstieg, D. & Pfurtscheller, G. (2010). An application framework for controlling an avatar in a desktop-based virtual environment via a software SSVEP brain–computer interface. *Presence: teleoperators and virtual environments*, **19**, 25–34.

Farwell, L.A. & Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, **70**, 510–523.

Flaherty-Craig, C., Eslinger, P., Stephens, B. & Simmons, Z. (2006). A rapid screening battery to identify frontal dysfunction in patients with ALS. *Neurology*, **67**, 2070–2072.

Ford, J.M., Roth, W.T. & Kopell, B.S. (1976). Attention effects on auditory evoked potentials to infrequent events. *Biological Psychology*, **4**, 65–77.

Fullmer, H.M., Siedler, H.D., Krooth, R.S. & Kurland, L.T. (1960). A cutaneous disorder of connective tissue in amyotrophic lateral sclerosis: a histochemical study. *Neurology*, **10**, 717–717.

Furdea, A., Halder, S., Krusienski, D.J., Bross, D., Nijboer, F., Birbaumer, N. & Kübler, A. (2009). An auditory oddball (P300) spelling system for brain-computer interfaces. *Psychophysiology*, **46**, 617–625.

General Assembly of the World Medical Association (2014). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *The Journal of the American College of Dentists*, **81**, 14–18.

Georgesco, M., Salerno, A. & Camu, W. (1997). Somatosensory evoked potentials elicited by stimulation of lower-limb nerves in amyotrophic lateral sclerosis. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, **104**, 333–342.

Gerven, M.v., Farquhar, J., Schaefer, R., Vlek, R., Geuze, J., Nijholt, A., Ramsey, N., Haselager, P., Vuurpijl, L., Gielen, S. & Desain, P. (2009). The brain–computer interface cycle. *Journal of Neural Engineering*, **6**, 041001.

Gescheider, G.A., Bolanowski, S.J., Hall, K.L., Hoffman, K.E. & Verrillo, R.T. (1994). The Effects of Aging on Information-Processing Channels in the Sense of Touch: I. Absolute Sensitivity. *Somatosensory & Motor Research*, **11**, 345–357.

Grier, R.A. (2015). How High is High? A Meta-Analysis of NASA-TLX Global Workload Scores. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, **59**, 1727–1731.

Grosse-Wentrup, M. & Schölkopf, B. (2013). A review of performance variations in SMR-based Brain- Computer interfaces (BCIs). In *Brain-Computer Interface Research*, 39–51, Springer.

Grossman, M. (2019). Amyotrophic lateral sclerosis—a multisystem neurodegenerative disorder. *Nature Reviews Neurology*, **15**, 5.

Guger, C., Edlinger, G., Harkam, W., Niedermayer, I. & Pfurtscheller, G. (2003). How many people are able to operate an EEG-based brain-computer interface (BCI)? *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, **11**, 145–147.

Halata, Z. (1988). Chapter 24 Ruffini corpuscle – a stretch receptor in the connective tissue of the skin and locomotion apparatus. In W. Hamann & A. Iggo, eds., *Progress in Brain Research*, vol. 74 of *Transduction and Cellular Mechanisms in Sensory Receptors*, 221–229, Elsevier.

Halata, Z. & Munger, B.L. (1981). Identification of the Ruffini corpuscle in human hairy skin. *Cell and Tissue Research*, **219**, 437–440.

Halata, Z., Grim, M. & Bauman, K.I. (2003). Friedrich Sigmund Merkel and his "Merkel cell", morphology, development, and physiology: review and new results. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology: An Official Publication of the American Association of Anatomists*, **271**, 225–239.

Halder, S., Hammer, E.M., Kleih, S.C., Bogdan, M., Rosenstiel, W., Birbaumer, N. & Kübler, A. (2013). Prediction of auditory and visual p300 brain-computer interface aptitude. *PloS one*, **8**.

Halder, S., Käthner, I. & Kübler, A. (2016). Training leads to increased auditory brain–computer interface performance of end-users with motor impairments. *Clinical Neurophysiology*, **127**, 1288–1296.

Halder, S., Raederscheidt, J., Hess, R., Eck, D., Schilling, K. & Kübler, A. (2017). Tactile brain-Computer Interface control of a Mobile Platform in a Real World Environment using a low-Cost Electroencephalography Headset. In *Proceedings of the 7th Graz Brain-Computer Interface Conference*, 171–175, Graz.

Halder, S., Takano, K. & Kansaku, K. (2018). Comparison of Four Control Methods for a Five-Choice Assistive Technology. *Frontiers in Human Neuroscience*, **12**.

Halder, S., Leinfelder, T., Schulz, S.M. & Kübler, A. (2019). Neural mechanisms of training an auditory event-related potential task in a brain–computer interface context. *Human Brain Mapping*, **40**, 2399–2412.

Hamilton, R.H. & Pascual-Leone, A. (1998). Cortical plasticity associated with Braille learning. *Trends in cognitive sciences*, **2**, 168–174.

Hammad, M., Silva, A., Glass, J., Sladky, J.T. & Benatar, M. (2007). Clinical, electrophysiologic, and pathologic evidence for sensory abnormalities in ALS. *Neurology*, **69**, 2236–2242.

Hammer, E.M., Halder, S., Blankertz, B., Sannelli, C., Dickhaus, T., Kleih, S., Müller, K.R. & Kübler, A. (2012). Psychological predictors of SMR-BCI performance. *Biological Psychology*, **89**, 80–86.

Hammer, E.M., Kaufmann, T., Kleih, S.C., Blankertz, B. & Kübler, A. (2014). Visuo-motor coordination ability predicts performance with brain-computer interfaces controlled by modulation of sensorimotor rhythms (SMR). *Frontiers in human neuroscience*, **8**, 574.

Hanagasi, H.A., Gurvit, I.H., Ermutlu, N., Kaptanoglu, G., Karamursel, S., Idrisoglu, H.A., Emre, M. & Demiralp, T. (2002). Cognitive impairment in amyotrophic lateral sclerosis: evidence from neuropsychological investigation and event-related potentials. *Cognitive Brain Research*, **14**, 234–244.

Hardiman, O., Al-Chalabi, A., Brayne, C., Beghi, E., Berg, L.H.v.d., Chio, A., Martin, S., Logroscino, G. & Rooney, J. (2017). The changing picture of amyotrophic lateral sclerosis: lessons from European registers. *J Neurol Neurosurg Psychiatry*, **88**, 557–563.

Hart, S.G. & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P.A. Hancock & N. Meshkati, eds., *Advances in Psychology*, vol. 52 of *Human Mental Workload*, 139–183, North-Holland.

Heath, P.R. & Shaw, P.J. (2002). Update on the glutamatergic neurotransmitter system and the role of excitotoxicity in amyotrophic lateral sclerosis. *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, **26**, 438–458.

Hecht, M., Hillemacher, T., Gräsel, E., Tigges, S., Winterholler, M., Heuss, D., Hilz, M.J. & Neundörfer, B. (2002). Subjective experience and coping in ALS. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders*, **3**, 225–231.

Herweg, A., Gutzeit, J., Kleih, S. & Kübler, A. (2016). Wheelchair control by elderly participants in a virtual environment with a brain-computer interface (BCI) and tactile stimulation. *Biological psychology*, **121**, 117–124.

Höhne, J., Schreuder, M., Blankertz, B. & Tangermann, M. (2011). A novel 9-class auditory ERP paradigm driving a predictive text entry system. *Frontiers in neuroscience*, **5**, 99.

Hill, N.J. & Schölkopf, B. (2012). An online brain–computer interface based on shifting attention to concurrent streams of auditory stimuli. *Journal of neural engineering*, **9**, 026011.

Hitziger, S. (2015). *Modeling the variability of electrical activity in the brain*. Ph.D. thesis, Université Nice Sophia Antipolis.

Hämäläinen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J. & Lounasmaa, O.V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of modern Physics*, **65**, 413.

Holz, E.M., Botrel, L., Kaufmann, T. & Kübler, A. (2015a). Long-Term Independent Brain-Computer Interface Home Use Improves Quality of Life of a Patient in the Locked-In State: A Case Study. *Archives of Physical Medicine and Rehabilitation*, **96**, S16–S26.

Holz, E.M., Botrel, L. & Kübler, A. (2015b). Independent home use of Brain Painting improves quality of life of two artists in the locked-in state diagnosed with amyotrophic lateral sclerosis. *Brain-Computer Interfaces*, **2**, 117–134.

Huettel, S.A. & McCarthy, G. (2004). What is odd in the oddball task?: Prefrontal cortex is activated by dynamic changes in response strategy. *Neuropsychologia*, **42**, 379–386.

Hwang, H.J., Ferreria, V.Y., Ulrich, D., Kilic, T., Chatziliadis, X., Blankertz, B. & Treder, M. (2015). A gaze independent brain-computer interface based on visual stimulation through closed eyelids. *Scientific reports*, **5**, 15890.

Iglesias, C., Sangari, S., Mendili, M.M.E., Benali, H., Marchand-Pauvert, V. & Pradat, P.F. (2015). Electrophysiological and spinal imaging evidences for sensory dysfunction in amyotrophic lateral sclerosis. *BMJ Open*, **5**.

Imai, T., Kamping, S., Breitenstein, C., Pantev, C., Lütkenhöner, B. & Knecht, S. (2003). Learning of tactile frequency discrimination in humans. *Human Brain Mapping*, **18**, 260–271.

Isak, B., Tankisi, H., Johnsen, B., Pugdahl, K., Torvin MØLler, A., Finnerup, N.B., Christensen, P.B. & Fuglsang-Frederiksen, A. (2016). Involvement of distal sensory nerves in amyotrophic lateral sclerosis. *Muscle & Nerve*, **54**, 1086–1092.

ISO-9241-210 (2010). Ergonomics of human-system interaction. *Human-centred design for interactive systems*.

Iversen, I.H., Ghanayim, N., Kübler, A., Neumann, N., Birbaumer, N. & J. Kaiser (2008). A brain–computer interface tool to assess cognitive functions in completely paralyzed patients with amyotrophic lateral sclerosis. *Clinical Neurophysiology*, **119**, 2214–2223.

Iwasaki, T., Goto, N., Goto, J., Ezure, H. & Moriyama, H. (2003). The aging of human Meissner's corpuscles as evidenced by parallel sectioning. *Okajimas folia anatomica Japonica*, **79**, 185–189.

Jamal, G.A., Weir, A.I., Hansen, S. & Ballantyne, J.P. (1985). Sensory involvement in motor neuron disease: further evidence from automated thermal threshold determination. *Journal of Neurology, Neurosurgery & Psychiatry*, **48**, 906–910.

Jasper, H.H. (1958). The 10/20 international electrode system. *EEG and Clinical Neurophysiology*, **10**, 370–375.

Kanai, K., Kuwabara, S., Misawa, S., Tamura, N., Ogawara, K., Nakata, M., Sawai, S., Hattori, T. & Bostock, H. (2006). Altered axonal excitability properties in amyotrophic lateral sclerosis: impaired potassium channel function related to disease stage. *Brain*, **129**, 953–962.

Kandel, E.R., Schwartz, J.H. & Jessel, T.M. (2000). *Principles of neural science*. McGraw-Hill, Health Professions Division, New York, 5th edn.

Kaongoen, N. & Jo, S. (2018). An auditory P300-based brain-computer interface using Ear-EEG. In *2018 6th International Conference on Brain-Computer Interface (BCI)*, 1–4, IEEE.

Kaufmann, T., Schulz, S.M., Grünzinger, C. & Kübler, A. (2011). Flashing characters with famous faces improves ERP-based brain–computer interface performance. *Journal of Neural Engineering*, **8**, 056016.

Kaufmann, T., Holz, E.M. & Kübler, A. (2013). Comparison of tactile, auditory, and visual modality for brain-computer interface use: a case study with a patient in the locked-in state. *Frontiers in Neuroscience*, **7**.

Kaufmann, T., Herweg, A. & Kübler, A. (2014). Toward brain-computer interface based wheelchair control utilizing tactually-evoked event-related potentials. *Journal of NeuroEngineering and Rehabilitation*, **11**, 7.

Kübler, A. (2013). Brain-computer interfacing: science fiction has come true. *Brain*, **136**, 2001–2004.

Kübler, A. (2019). The history of BCI: From a vision for the future to real support for personhood in people with locked-in syndrome. *Neuroethics*, 1–18.

Kübler, A., Kotchoubey, B., Hinterberger, T., Ghanayim, N., Perelmouter, J., Schauer, M., Fritsch, C., Taub, E. & Birbaumer, N. (1999). The thought translation device: a neurophysiological approach to communication in total motor paralysis. *Experimental Brain Research*, **124**, 223–232.

Kübler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J.R. & Birbaumer, N. (2001a). Brain–computer communication: Unlocking the locked in. *Psychological bulletin*, **127**, 358.

Kübler, A., Neumann, N., Kaiser, J., Kotchoubey, B., Hinterberger, T. & Birbaumer, N.P. (2001b). Brain-computer communication: Self-regulation of slow cortical potentials for verbal communication. *Archives of Physical Medicine and Rehabilitation*, **82**, 1533–1539.

Kübler, A., Neumann, N., Wilhelm, B., Hinterberger, T. & Birbaumer, N. (2004). Predictability of brain-computer communication. *Journal of Psychophysiology*, **18**, 121–129.

Kübler, A., Nijboer, F., Mellinger, J., Vaughan, T.M., Pawelzik, H., Schalk, G., McFarland, D.J., Birbaumer, N. & Wolpaw, J.R. (2005). Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, **64**, 1775–1777.

Kübler, A., Dixon, V. & Garavan, H. (2006). Automaticity and Reestablishment of Executive Control—An fMRI Study. *Journal of Cognitive Neuroscience*, **18**, 1331–1342.

Kübler, A., Furdea, A., Halder, S., Hammer, E.M., Nijboer, F. & Kotchoubey, B. (2009). A brain-computer interface controlled auditory event-related potential (P300) spelling system for locked-in patients. *Annals of the New York Academy of Sciences*, **1157**, 90–100.

Kübler, A., Blankertz, B., Müller, K.R. & Neuper, C. (2011). A model of BCI-control. In *Proceedings of the 5th Graz Brain-Computer Interface Conference*, 100–103, Graz.

Kübler, A., Mattia, D., Rupp, R. & Tangermann, M. (2013). Facing the challenge: Bringing brain–computer interfaces to end-users. *Artificial Intelligence in Medicine*, **59**, 55–60.

Kübler, A., Holz, E.M., Riccio, A., Zickler, C., Kaufmann, T., Kleih, S.C., Staiger-Sälzer, P., Desideri, L., Hoogerwerf, E.J. & Mattia, D. (2014). The User-Centered Design as Novel Perspective for Evaluating the Usability of BCI-Controlled Applications. *PLOS ONE*, **9**, e112392.

Kübler, A., Müller-Putz, G. & Mattia, D. (2015). User-centred design in brain–computer interface research and development. *Annals of Physical and Rehabilitation Medicine*, **58**, 312–314.

Kelly, A.M.C. & Garavan, H. (2005). Human Functional Neuroimaging of Brain Changes Associated with Practice. *Cerebral Cortex*, **15**, 1089–1102, publisher: Oxford Academic.

Kepp, K.P. (2019). A quantitative model of human neurodegenerative diseases involving protein aggregation. *Neurobiology of aging*, **80**, 46–55.

Kida, T., Nishihira, Y., Hatta, A., Wasaka, T., Tazoe, T., Sakajiri, Y., Nakata, H., Kaneda, T., Kuroiwa, K., Akiyama, S., Sakamoto, M., Kamijo, K. & Higashiura, T. (2004). Resource allocation and somatosensory P300 amplitude during dual task: effects of tracking speed and predictability of tracking direction. *Clinical Neurophysiology*, **115**, 2616–2628.

Kida, T., Kaneda, T. & Nishihira, Y. (2012). Dual-task repetition alters event-related brain potentials and task performance. *Clinical Neurophysiology*, **123**, 1123–1130.

Kiernan, M.C., Vucic, S., Cheah, B.C., Turner, M.R., Eisen, A., Hardiman, O., Burrell, J.R. & Zoing, M.C. (2011). Amyotrophic lateral sclerosis. *The Lancet*, **377**, 942–955.

Kleih, S.C. & Kübler, A. (2015). Psychological factors influencing brain-computer interface (BCI) performance. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 3192–3196, IEEE.

Kleih, S.C., Nijboer, F., Halder, S. & Kübler, A. (2010). Motivation modulates the P300 amplitude during brain–computer interface use. *Clinical Neurophysiology*, **121**, 1023–1031.

Kleih, S.C., Kaufmann, T., Zickler, C., Halder, S., Leotta, F., Cincotti, F., Aloise, F., Riccio, A., Herbert, C., Mattia, D. & Kübler, A. (2011). Chapter

2 - Out of the frying pan into the fire—the P300-based BCI faces real-world challenges. In J. Schouenborg, M. Garwicz & N. Danielsen, eds., *Progress in Brain Research*, vol. 194 of *Brain Machine Interfaces: Implications for Science, Clinical Practice and Society*, 27–46, Elsevier.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, **38**, 557–577.

Škola, F., Tinková, S. & Liarokapis, F. (2019). Progressive training for motor imagery brain-computer interfaces using gamification and virtual reality embodiment. *Frontiers in Human Neuroscience*, **13**, 329.

Kraus, N., McGee, T., Carrell, T.D., King, C., Tremblay, K. & Nicol, T. (1995). Central auditory system plasticity associated with speech discrimination training. *Journal of cognitive neuroscience*, **7**, 25–32.

Krusienski, D.J., Sellers, E.W., Cabestaing, F., Bayoudh, S., McFarland, D.J., Vaughan, T.M. & Wolpaw, J.R. (2006). A comparison of classification techniques for the P300 Speller. *Journal of neural engineering*, **3**, 299.

Krusienski, D.J., Sellers, E.W., McFarland, D.J., Vaughan, T.M. & Wolpaw, J.R. (2008). Toward enhanced P300 speller performance. *Journal of Neuroscience Methods*, **167**, 15–21.

Käthner, I., Ruf, C.A., Pasqualotto, E., Braun, C., Birbaumer, N. & Halder, S. (2013). A portable auditory P300 brain–computer interface with directional cues. *Clinical Neurophysiology*, **124**, 327–338.

Käthner, I., Wriessnegger, S.C., Müller-Putz, G.R., Kübler, A. & Halder, S. (2014). Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface. *Biological Psychology*, **102**, 118–129.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*, **4**, 863.

Lakerveld, J., Kotchoubey, B. & Kübler, A. (2008). Cognitive function in patients with late stage amyotrophic lateral sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, **79**, 25–29.

Laureys, S., Pellas, F., Van Eeckhout, P., Ghorbel, S., Schnakers, C., Perrin, F., Berré, J., Faymonville, M.E., Pantke, K.H., Damas, F., Lamy, M., Moonen, G. & Goldman, S. (2005). The locked-in syndrome : what is it like to be conscious but paralyzed and voiceless? In S. Laureys, ed., *Progress in Brain Research*, vol. 150 of *The Boundaries of Consciousness: Neurobiology and Neuropathology*, 495–611, Elsevier.

Logan, G.D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology/Revue canadienne de psychologie*, **39**, 367–386.

Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492–527.

Logroscino, G., Traynor, B.J., Hardiman, O., Chiò, A., Mitchell, D., Swingler, R.J., Millul, A., Benn, E., Beghi, E. & Eurals, F. (2010). Incidence of amyotrophic lateral sclerosis in Europe. *Journal of Neurology, Neurosurgery & Psychiatry*, **81**, 385–390.

Lorenz, R., Pascual, J., Blankertz, B. & Vidaurre, C. (2014). Towards a holistic assessment of the user experience with hybrid BCIs. *Journal of neural engineering*, **11**, 035007.

Lotte, F., Larrue, F. & Mühl, C. (2013). Flaws in current human training protocols for spontaneous brain-computer interfaces: lessons learned from instructional design. *Frontiers in human neuroscience*, **7**, 568.

Luck, S.J. (2005). *The event-related potential technique*. Cambridge, MA: MIT Press.

Luck, S.J. (2012). Event-related potentials. In H. Cooper, P.M. Camic, D.L. Long, A.T. Panter, D. Rindskopf & K.J. Sher, eds., *APA handbook of research methods in psychology, Vol 1: Foundations, planning, measures, and psychometrics*, 523–546, American Psychological Association, Washington, DC, US.

Ludolph, A.C., Langen, K.J., Regard, M., Herzog, H., Kemper, B., Kuwert, T., Böttger, I.G. & Feinendegen, L. (1992). Frontal lobe function in amyotrophic lateral sclerosis: a neuropsychologic and positron emission tomography study. *Acta Neurologica Scandinavica*, **85**, 81–89.

Mahns, D.A., Perkins, N.M., Sahai, V., Robinson, L. & Rowe, M.J. (2006). Vibrotactile Frequency Discrimination in Human Hairy Skin. *Journal of Neurophysiology*, **95**, 1442–1450.

Mayaud, L., Cabanilles, S. & Azabou, E. (2016). Analysis of Patient Need for Brain–Computer Interfaces. *Brain–Computer Interfaces 2: Technology and Applications*, 101–121, publisher: Wiley Online Library.

McCane, L.M., Sellers, E.W., McFarland, D.J., Mak, J.N., Carmack, C.S., Zeitlin, D., Wolpaw, J.R. & Vaughan, T.M. (2014). Brain-computer interface (BCI) evaluation in people with amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis and frontotemporal degeneration*, **15**, 207–215.

McCane, L.M., Heckman, S.M., McFarland, D.J., Townsend, G., Mak, J.N., Sellers, E.W., Zeitlin, D., Tenteromano, L.M., Wolpaw, J.R. & Vaughan, T.M.

(2015). P300-based brain-computer interface (BCI) event-related potentials (ERPs): People with amyotrophic lateral sclerosis (ALS) vs. age-matched controls. *Clinical Neurophysiology*, **126**, 2124–2131.

McRobbie, D.W., Moore, E.A., Graves, M.J. & Prince, M.R. (2017). *MRI from Picture to Proton*. Cambridge university press.

Miller, R.G., Mitchell, J.D. & Moore, D.H. (2012). Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database of Systematic Reviews*.

Mitchell, J. & Borasio, G. (2007). Amyotrophic lateral sclerosis. *The Lancet*, **369**, 2031–2041.

Müller-Putz, G., Scherer, R., Brunner, C., Leeb, R. & Pfurtscheller, G. (2008). Better than random: a closer look on BCI results. *International Journal of Bioelectromagnetism*, **10**, 52–55.

Münssinger, J.I., Halder, S., Kleih, S.C., Furdea, A., Raco, V., Hösle, A. & Kubler, A. (2010). Brain painting: first evaluation of a new brain–computer interface application with ALS-patients and healthy volunteers. *Frontiers in neuroscience*, **4**, 182.

Nagarajan, S.S., Blake, D.T., Wright, B.A., Byl, N. & Merzenich, M.M. (1998). Practice-Related Improvements in Somatosensory Interval Discrimination Are Temporally Specific But Generalize across Skin Location, Hemisphere, and Modality. *Journal of Neuroscience*, **18**, 1559–1570.

Nakagawa, S. (2004). A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral ecology*, **15**, 1044–1045.

Neuper, C. & Pfurtscheller, G. (1999). Motor imagery and ERD In: Event-Related Desynchronization. *Handbook of Electroencephalography and Clinical Neurophysiology. Elsevier, Amsterdam303–325*.

Neuper, C., Müller-Putz, G.R., Scherer, R. & Pfurtscheller, G. (2006). Motor imagery and EEG-based control of spelling devices and neuroprostheses. *Progress in Brain Research*, **159**, 393–409.

Niebrój-Dobosz, I., Dziewulska, D. & Kwieciński, H. (2004). Oxidative damage to proteins in the spinal cord in amyotrophic lateral sclerosis (ALS). *Folia neuropathologica*, **42**, 151–156.

Nijboer, F. (2015). Technology transfer of brain-computer interfaces as assistive technology: Barriers and opportunities. *Annals of Physical and Rehabilitation Medicine*, **58**, 35–38.

Nijboer, F., Furdea, A., Gunst, I., Mellinger, J., McFarland, D.J., Birbaumer, N. & Kübler, A. (2008a). An auditory brain–computer interface (BCI). *Journal of neuroscience methods*, **167**, 43–50.

NIJBOER, F., SELLERS, E.W., MELLINGER, J., JORDAN, M.A., MATUZ, T., FURDEA, A., HALDER, S., MOCHTY, U., KRUSIENSKI, D.J. & VAUGHAN, T.M. (2008b). A P300-based brain–computer interface for people with amyotrophic lateral sclerosis. *Clinical neurophysiology*, **119**, 1909–1916.

NIJBOER, F., BIRBAUMER, N. & KUBLER, A. (2010). The influence of psychological state and motivation on brain–computer interface performance in patients with amyotrophic lateral sclerosis–a longitudinal study. *Frontiers in neuroscience*, **4**, 55.

NUNEZ, P.L., WOLPAW, J. & WOLPAW, E.W. (2012). Electric and magnetic fields produced by the brain. *Brain–computer interfaces: principles and practice. Oxford University Press, New York*, 45–64.

OGAWA, S., LEE, T.M., NAYAK, A.S. & GLYNN, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic resonance in medicine*, **14**, 68–78.

ONISHI, A., TAKANO, K., KAWASE, T., ORA, H. & KANSAKU, K. (2017). Affective Stimuli for an Auditory P300 Brain-Computer Interface. *Frontiers in Neuroscience*, **11**.

ONO, S., MECHANIC, G.L. & YAMAUCHI, M. (1990). Amyotrophic lateral sclerosis: unusually low content of collagen in skin. *Journal of the Neurological Sciences*, **100**, 234–237.

ORTNER, R., RODRIGUEZ, J. & GUGER, C. (2012). A tactile P300 brain-computer interface. Tech. rep., technical report 3. Austria: g. tec Medical Engineering GmbH.

ORTNER, R., LUGO, Z., NOIRHOMME, Q., LAUREYS, S. & GUGER, C. (2014). A tactile Brain-Computer Interface for severely disabled patients. In *2014 IEEE Haptics Symposium (HAPTICS)*, 235–237, IEEE, Houston, TX, USA.

PASQUALOTTO, E., FEDERICI, S. & BELARDINELLI, M.O. (2012). Toward functioning and usable brain–computer interfaces (BCIs): A literature review. *Disability and Rehabilitation: Assistive Technology*, **7**, 89–103.

PATTERSON, J.R. & GRABOIS, M. (1986). Locked-in syndrome: a review of 139 cases. *Stroke*, **17**, 758–764.

PAULUS, K.S., MAGNANO, I., PIRAS, M.R., SOLINAS, M.A., SOLINAS, G., SAU, G.F. & AIELLO, I. (2002). Visual and auditory event-related potentials in sporadic amyotrophic lateral sclerosis. *Clinical neurophysiology*, **113**, 853–861.

PFURTSCHELLER, G. (1992). Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest. *Electroencephalography and clinical neurophysiology*, **83**, 62–69.

PFURTSCHELLER, G. & DA SILVA, F.L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology*, **110**, 1842–1857.

PICCININ, M.A. & SCHWARTZ, J. (2018). *Histology, Meissner Corpuscle*. StatPearls Publishing.

PLAITAKIS, A. & CAROSCIO, J.T. (1987). Abnormal glutamate metabolism in amyotrophic lateral sclerosis. *Annals of Neurology*, **22**, 575–579.

PLUM, F. & POSNER, J.B. (1966). The Diagnosis of Stupor and Coma. *Philadelphia, FA Davis Co*, 18–19.

POKORNY, C., BREITWIESER, C. & MÜLLER-PUTZ, G.R. (2016). The Role of Transient Target Stimuli in a Steady-State Somatosensory Evoked Potential-Based Brain–Computer Interface Setup. *Frontiers in Neuroscience*, **10**.

POLDRACK, R.A. & GABRIELI, J.D. (2001). Characterizing the neural mechanisms of skill learning and repetition priming: evidence from mirror reading. *Brain*, **124**, 67–82.

POLICH, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, **118**, 2128–2148.

POLICH, J. & KOK, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*, **41**, 103–146.

PUTNAM, L.E. & ROTH, W.T. (1990). Effects of stimulus repetition, duration, and rise time on startle blink and automatically elicited P300. *Psychophysiology*, **27**, 275–297.

RAMSEY, N.F. (2012). Signals reflecting brain metabolic activity. *Brain–computer interfaces: principles and practice. Oxford University Press, New York*, 65–80.

RAPPAPORT, M., CLIFFORD, J.O. & WINTERFIELD, K.M. (1990). P300 response under active and passive attentional states and uni- and bimodality stimulus presentation conditions. *The Journal of neuropsychiatry and clinical neurosciences*, **2**, 399–407.

RECANZONE, G.H., MERZENICH, M.M., JENKINS, W.M., GRAJSKI, K.A. & DINSE, H.R. (1992). Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task. *journal of Neurophysiology*, **67**, 1031–1056.

REUTER, E.M., VOELCKER-REHAGE, C., VIELUF, S. & GODDE, B. (2014). Effects of age and expertise on tactile learning in humans. *European Journal of Neuroscience*, **40**, 2589–2599.

RICCIO, A., MATTIA, D., SIMIONE, L., OLIVETTI, M. & CINCOTTI, F. (2012). Eye-gaze independent EEG-based brain–computer interfaces for communication. *Journal of Neural Engineering*, **9**, 045001.

RICCIO, A., HOLZ, E.M., ARICÒ, P., LEOTTA, F., ALOISE, F., DESIDERI, L., RIMONDINI, M., KÜBLER, A., MATTIA, D. & CINCOTTI, F. (2015). Hybrid P300-Based Brain-Computer

Interface to Improve Usability for People With Severe Motor Disability: Electromyographic Signals for Error Correction During a Spelling Task. *Archives of Physical Medicine and Rehabilitation*, **96**, S54–S61.

Rooney, J., Byrne, S., Heverin, M., Corr, B., Elamin, M., Staines, A., Goldacre, B. & Hardiman, O. (2013). Survival Analysis of Irish Amyotrophic Lateral Sclerosis Patients Diagnosed from 1995–2010. *PLOS ONE*, **8**, e74733.

Roth, W.T., Blowers, G.H., Doyle, C.M. & Kopell, B.S. (1982). Auditory stimulus intensity effects on components of the late positive complex. *Electroencephalography and Clinical Neurophysiology*, **54**, 132–146.

Roth, W.T., Dorato, K.H. & Kopell, B.S. (1984). Intensity and task effects on evoked physiological responses to noise bursts. *Psychophysiology*, **21**, 466–481.

Rothstein, J.D., Martin, L.J. & Kuncl, R.W. (1992). Decreased Glutamate Transport by the Brain and Spinal Cord in Amyotrophic Lateral Sclerosis. *New England Journal of Medicine*, **326**, 1464–1468.

Sangari, S., Giron, A., Marrelec, G., Pradat, P.F. & Marchand-Pauvert, V. (2018). Abnormal cortical brain integration of somatosensory afferents in ALS. *Clinical Neurophysiology*, **129**, 874–884.

Sasaki, S. & Iwata, M. (2007). Mitochondrial alterations in the spinal cord of patients with sporadic amyotrophic lateral sclerosis. *Journal of Neuropathology & Experimental Neurology*, **66**, 10–16.

Schalk, G. & Mellinger, J. (2010). *A Practical Guide to Brain–Computer Interfacing with BCI2000: General-Purpose Software for Brain-Computer Interface Research, Data Acquisition, Stimulus Presentation, and Brain Monitoring*. Springer Science & Business Media.

Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N. & Wolpaw, J.R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, **51**, 1034–1043.

Schreuder, M., Rost, T. & Tangermann, M. (2011). Listen, you are writing! Speeding up online spelling with a dynamic auditory BCI. *Frontiers in neuroscience*, **5**, 112.

Schuster, C. (2018). Entwicklung und Evaluation eines taktil stimulierenden Brain-Computer-Interfaces und dessen Einsatz am Patienten (Unpublished Master Thesis), Technische Hochschule Nürnberg Georg Simon Ohm.

Sellers, E.W. & Donchin, E. (2006). A P300-based brain–computer interface: initial tests by ALS patients. *Clinical neurophysiology*, **117**, 538–548.

Sellers, E.W., Vaughan, T.M. & Wolpaw, J.R. (2010). A brain-computer interface for long-term independent home use. *Amyotrophic lateral sclerosis*, **11**, 449–455.

Seno, B., Matteucci, M. & Mainardi, L. (2010). The Utility Metric: A Novel Method to Assess the Overall Performance of Discrete Brain–Computer Interfaces. *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, **18**, 20–8.

Severens, M., Farquhar, J., Duysens, J. & Desain, P. (2013). A multi-signature brain–computer interface: use of transient and steady-state responses. *Journal of Neural Engineering*, **10**, 026005.

Severens, M., Van der Waal, M., Farquhar, J. & Desain, P. (2014). Comparing tactile and visual gaze-independent brain–computer interfaces in patients with amyotrophic lateral sclerosis and healthy users. *Clinical Neurophysiology*, **125**, 2297–2304.

Sharbrough, F. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *J clin Neurophysiol*, **8**, 200–202.

Silvoni, S., Konicar, L., Prats-Sedano, M.A., Garcia-Cossio, E., Genna, C., Volpato, C., Cavinato, M., Paggiaro, A., Veser, S., De Massari, D. & Birbaumer, N. (2016). Tactile event-related potentials in amyotrophic lateral sclerosis (ALS): Implications for brain-computer interface. *Clinical Neurophysiology*, **127**, 936–945.

Simon, N., Käthner, I., Ruf, C.A., Pasqualotto, E., Kübler, A. & Halder, S. (2015). An auditory multiclass brain-computer interface with natural stimuli: Usability evaluation with healthy participants and a motor impaired end user. *Frontiers in Human Neuroscience*, **8**.

Sitaram, R., Lee, S. & Birbaumer, N. (2012). 18| BCIS THAT USE BRAIN METABOLIC SIGNALS. *Brain-Computer Interfaces: Principles and Practice*, 301.

Smith, M.C. (1960). Nerve fibre degeneration in the brain in amyotrophic lateral sclerosis. *Journal of neurology, neurosurgery, and psychiatry*, **23**, 269.

Smulders, F.T., Kenemans, J.L., Schmidt, W.F. & Kok, A. (1999). Effects of task complexity in young and old adults: Reaction time and P300 latency are not always dissociated. *Psychophysiology*, **36**, 118–125.

Stippich, C. (2015). *Clinical functional MRI: presurgical functional neuroimaging*. Springer.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological)*, **36**, 111–133.

Sullivan, L.S., Klein, E., Brown, T., Sample, M., Pham, M., Tubig, P., Folland, R., Truitt, A. & Goering, S. (2018). Keeping disability in mind: a case study in implantable brain–computer Interface research. *Science and engineering ethics*, **24**, 479–504.

Sutton, S., Braren, M., Zubin, J. & John, E.R. (1965). Evoked-Potential Correlates of Stimulus Uncertainty. *Science*, **150**, 1187–1188.

Talbot, K. (2009). Motor neuron disease: The bare essentials. *Practical Neurology*, **9**, 303–309.

Thurlings, M.E., Brouwer, A.M., Van Erp, J.B., Blankertz, B. & Werkhoven, P.J. (2012a). Does bimodal stimulus presentation increase ERP components usable in BCIs? *Journal of neural engineering*, **9**, 045005.

Thurlings, M.E., van Erp, J.B., Brouwer, A.M., Blankertz, B. & Werkhoven, P. (2012b). Control-display mapping in brain–computer interfaces. *Ergonomics*, **55**, 564–580.

Thurlings, M.E., Van Erp, J.B.F., Brouwer, A.M. & Werkhoven, P. (2013). Controlling a Tactile ERP–BCI in a Dual Task. *IEEE Transactions on Computational Intelligence and AI in Games*, **5**, 129–140.

Treder, M.S. & Blankertz, B. (2010). (C) overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and brain functions*, **6**, 28.

Tremblay, K., Kraus, N., McGee, T., Ponton, C. & Otis, B. (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear and hearing*, **22**, 79–90.

van der Graaff, M.M., de Jong, J.M.B.V., Baas, F. & de Visser, M. (2009). Upper motor neuron and extra-motor neuron involvement in amyotrophic lateral sclerosis: A clinical and brain imaging review. *Neuromuscular Disorders*, **19**, 53–58.

Vaughan, T.M. (2003). Guest editorial brain-computer interface technology: a review of the second international meeting. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **11**, 94–109.

Velasco-Álvarez, F., Ron-Angevin, R., da Silva-Sauer, L., Sancha-Ros, S. & Blanca-Mena, M.J. (2011). Audio-Cued SMR Brain-Computer Interface to Drive a Virtual Wheelchair. In J. Cabestany, I. Rojas & G. Joya, eds., *Advances in Computational Intelligence*, Lecture Notes in Computer Science, 337–344, Springer Berlin Heidelberg.

Waal, M.v.d., Severens, M., Geuze, J. & Desain, P. (2012). Introducing the tactile speller: an ERP-based brain–computer interface for communication. *Journal of Neural Engineering*, **9**, 045002.

Walter, W.G. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *Nature*, **230**, 380–384.

Wickens, C.D., Heffley, E.F., Kramer, A.F. & Donchin, E. (1980). The Event-Related Brain Potential as an Index of Attention Allocation in Complex Displays. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, **24**, 297–301.

Wickremaratchi, M.M. & Llewelyn, J.G. (2006). Effects of ageing on touch. *Postgraduate Medical Journal*, **82**, 301–304.

Williams, C., Kozlowski, M.A., Hinton, D.R. & Miller, C.A. (1990). Degeneration of spinocerebellar neurons in amyotrophic lateral sclerosis. *Annals of Neurology*, **27**, 215–225.

Wolpaw, J. & Wolpaw, E.W. (2012). *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, USA, google-Books-ID: 0dST2Lg4KVYC.

Wolpaw, J.R., McFarland, D.J., Neat, G.W. & Forneris, C.A. (1991). An EEG-based brain-computer interface for cursor control. *Clinical Neurophysiology*, **78**, 252–259.

Zarei, S., Carr, K., Reiley, L., Diaz, K., Guerra, O., Altamirano, P.F., Pagani, W., Lodin, D., Orozco, G. & Chinea, A. (2015). A comprehensive review of amyotrophic lateral sclerosis. *Surgical Neurology International*, **6**.

Zickler, C., Riccio, A., Leotta, F., Hillian-Tress, S., Halder, S., Holz, E., Staiger-Sälzer, P., Hoogerwerf, E.J., Desideri, L., Mattia, D. & Kübler, A. (2011). A Brain-Computer Interface as Input Channel for a Standard Assistive Technology Software. *Clinical EEG and Neuroscience*, **42**, 236–244.

Ziebell, P., Stümpfig, J., Kleih, S., Kübler, A., Latoschik, M.E. & Halder, S. (2019). Introducing a Motivating Training Study Design to Compare Auditory and Tactile Streaming-Based P300 BCIs. In *Proceedings of the 8th Graz Brain-Computer Interface Conference*, 249–254, Graz.

Appendix A

Supplementary Figures

Study I

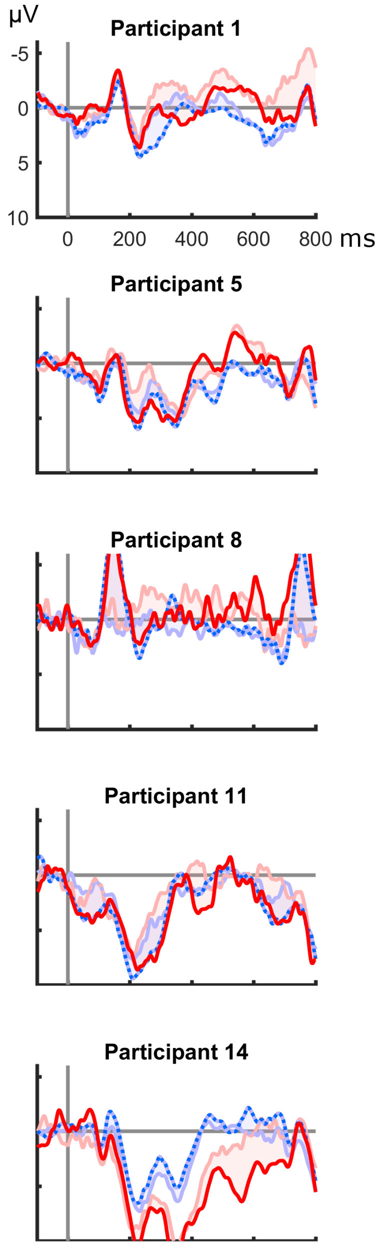

Figure A.1: Post-stimulus epochs at Fz.

Study I

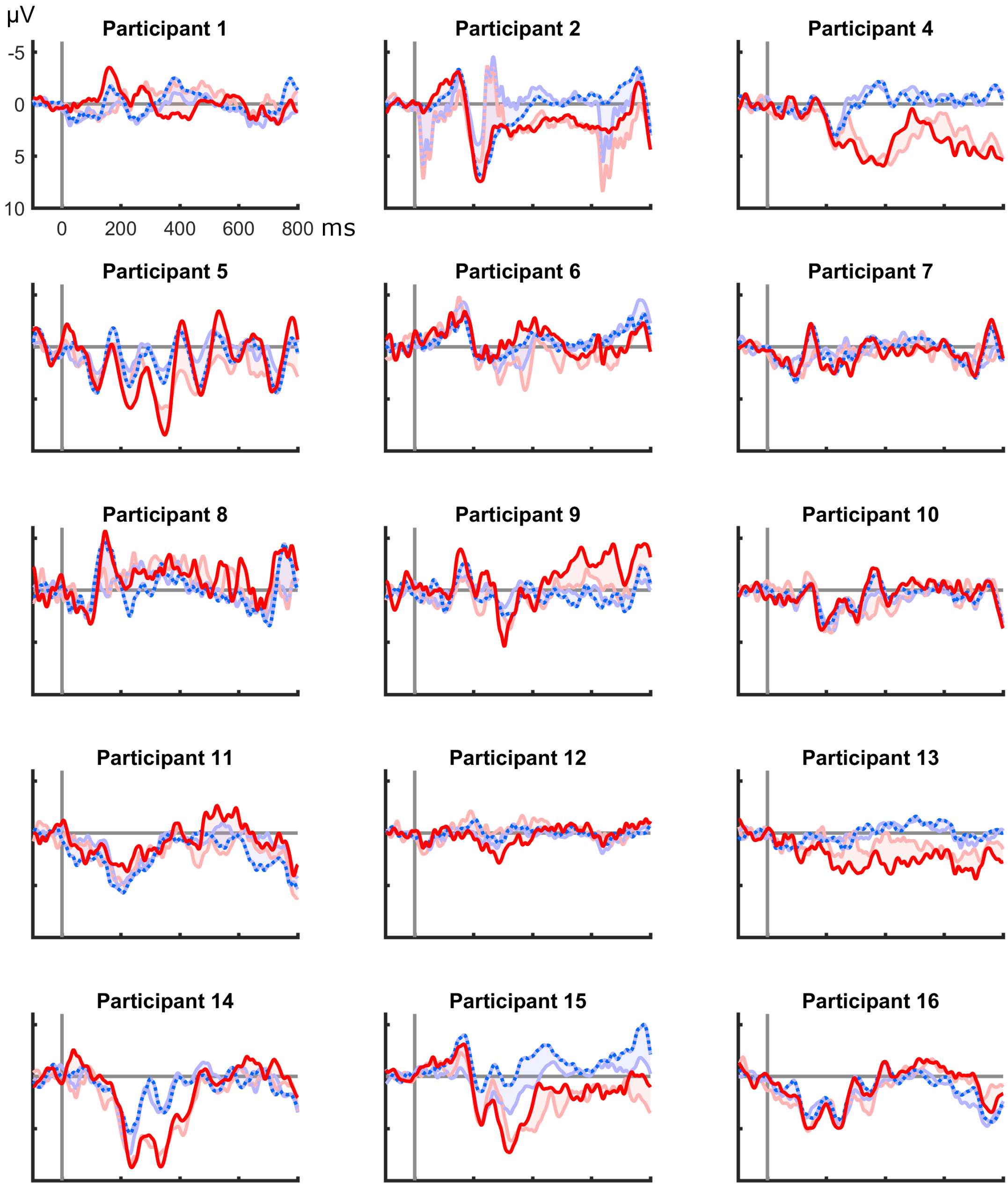

Figure A.2: Post-stimulus epochs at Pz.

Study II

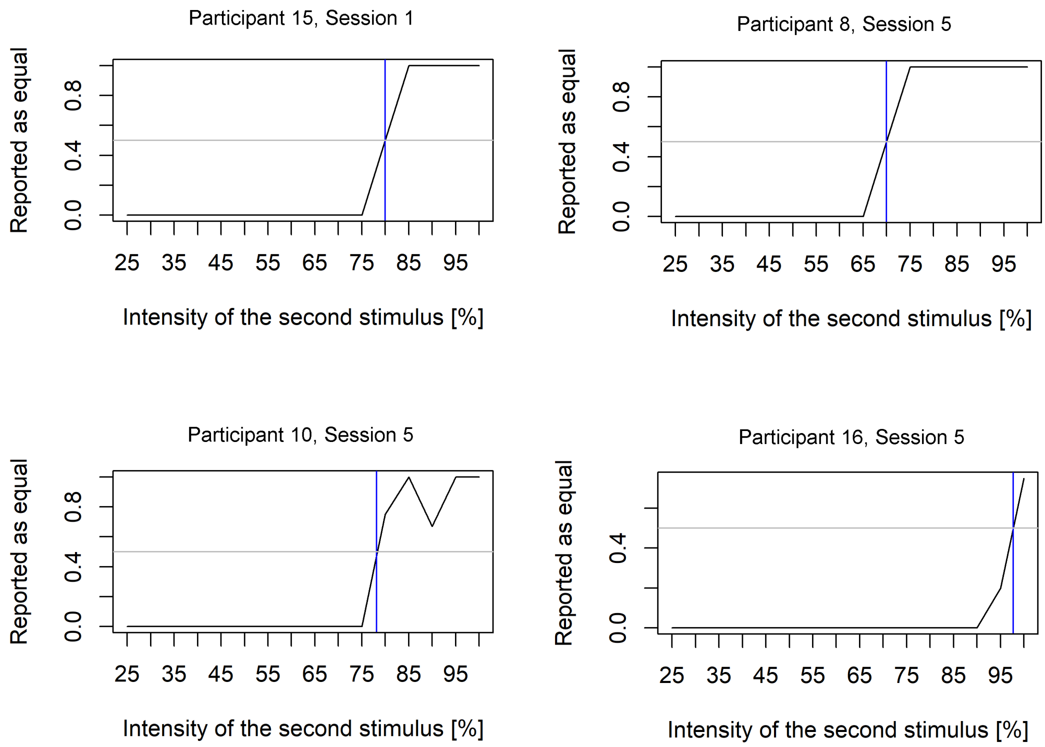

Figure A.3: Manually determined sensory thresholds

Study II

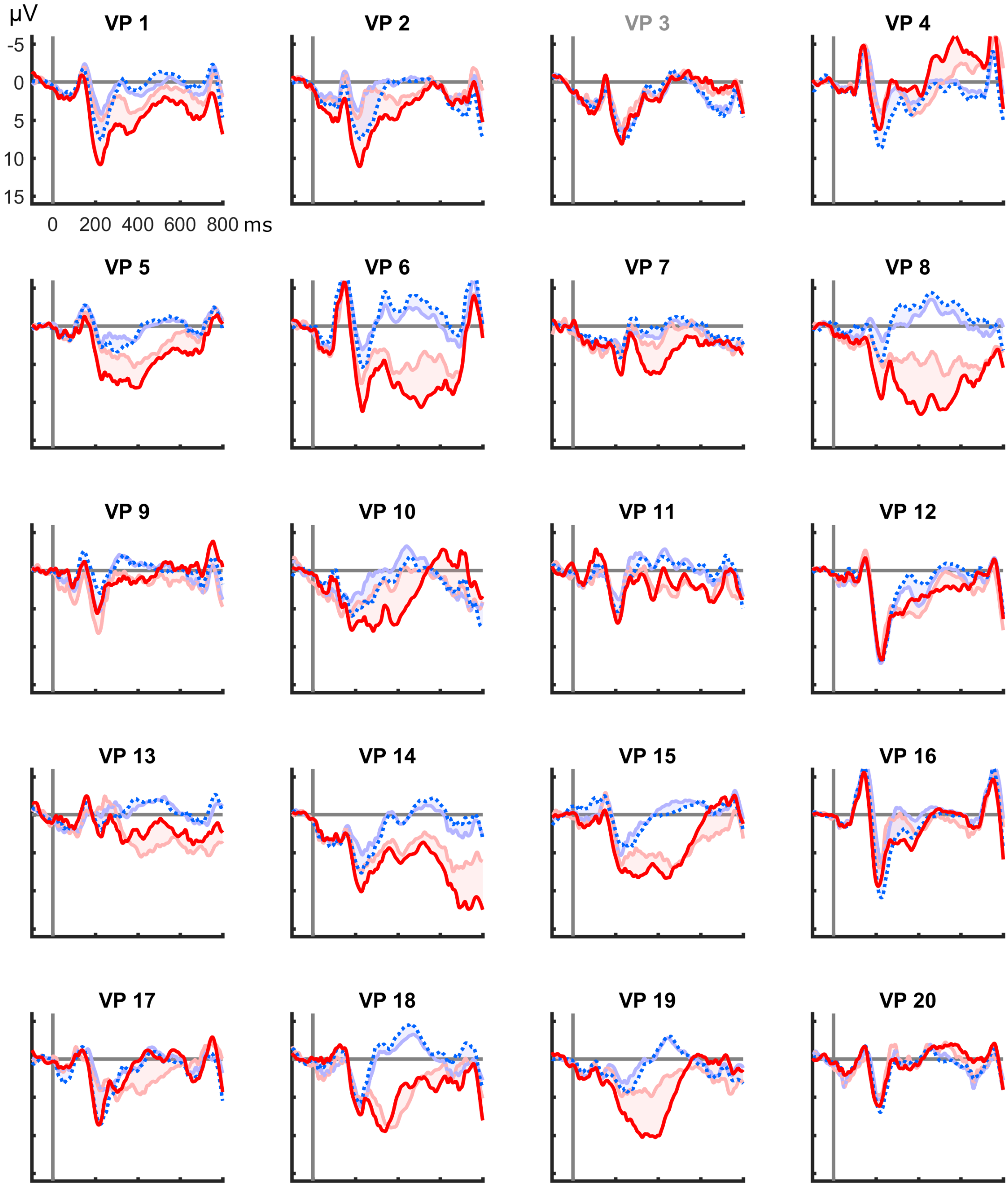

Figure A.4: Post-stimulus epochs at Cz.

Study II

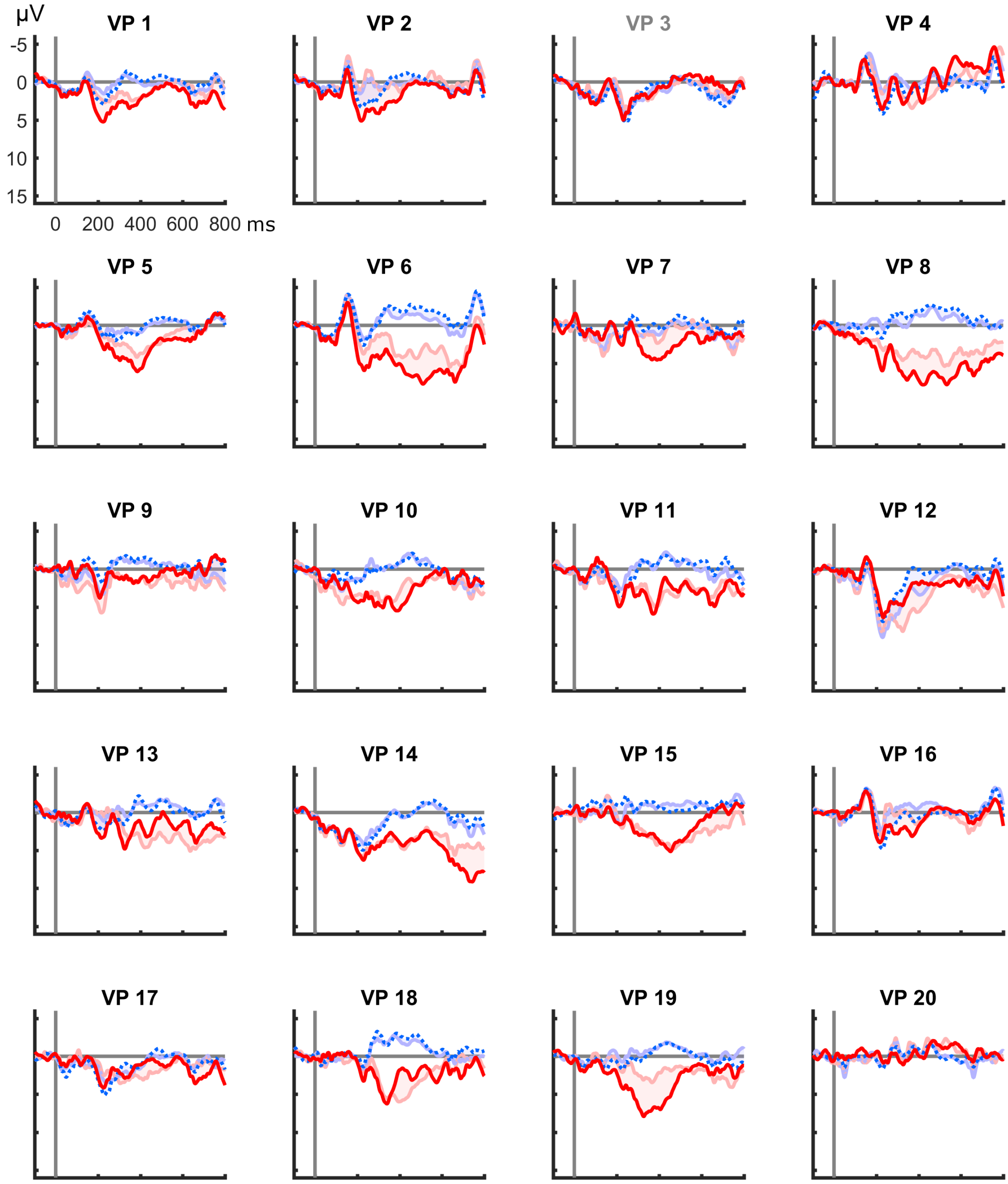

Figure A.5: Post-stimulus epochs at Pz.

Study III

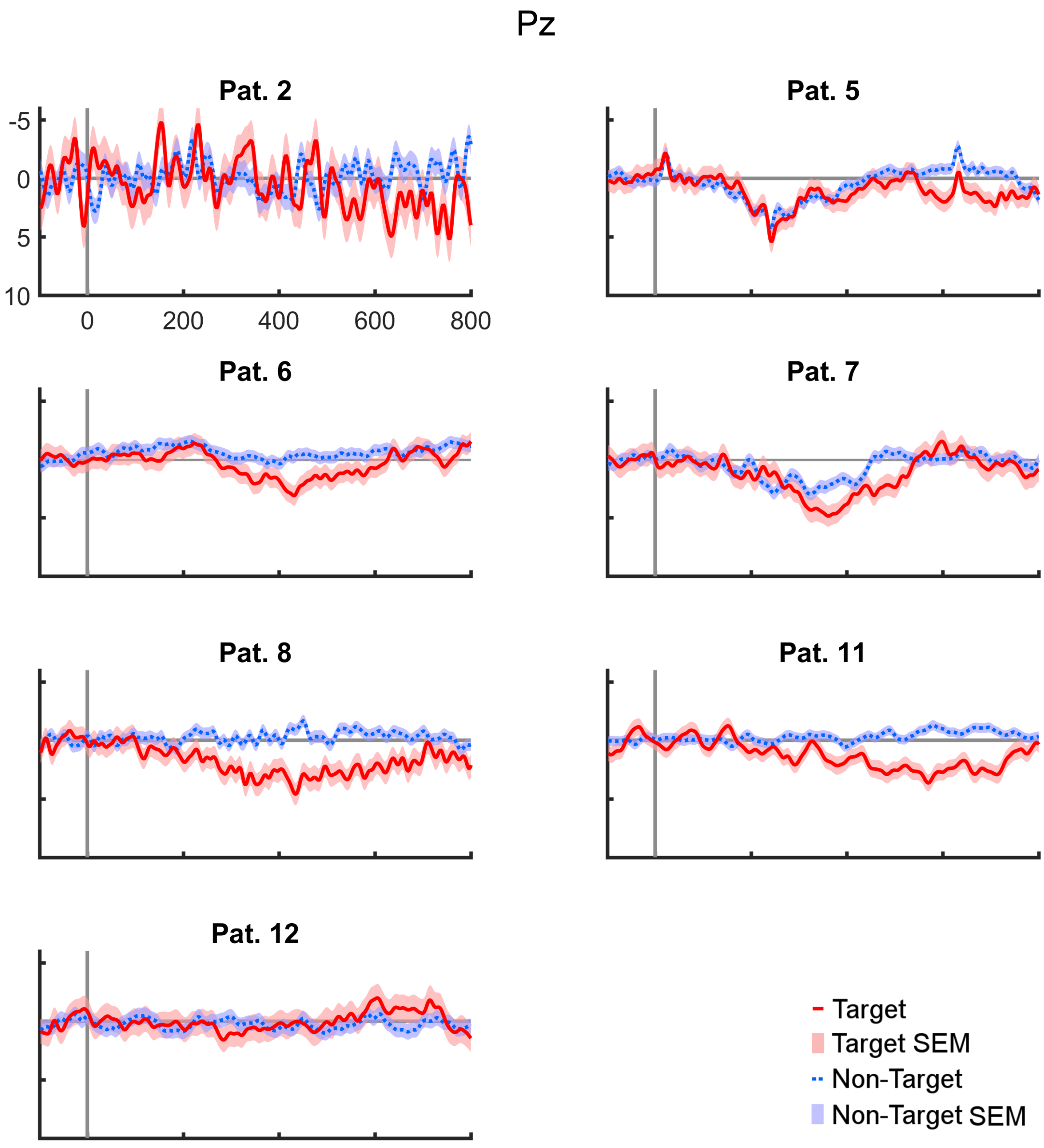

Figure A.6: Denmark: Post-stimulus epochs at Pz. Pat. = Patient.

Study III

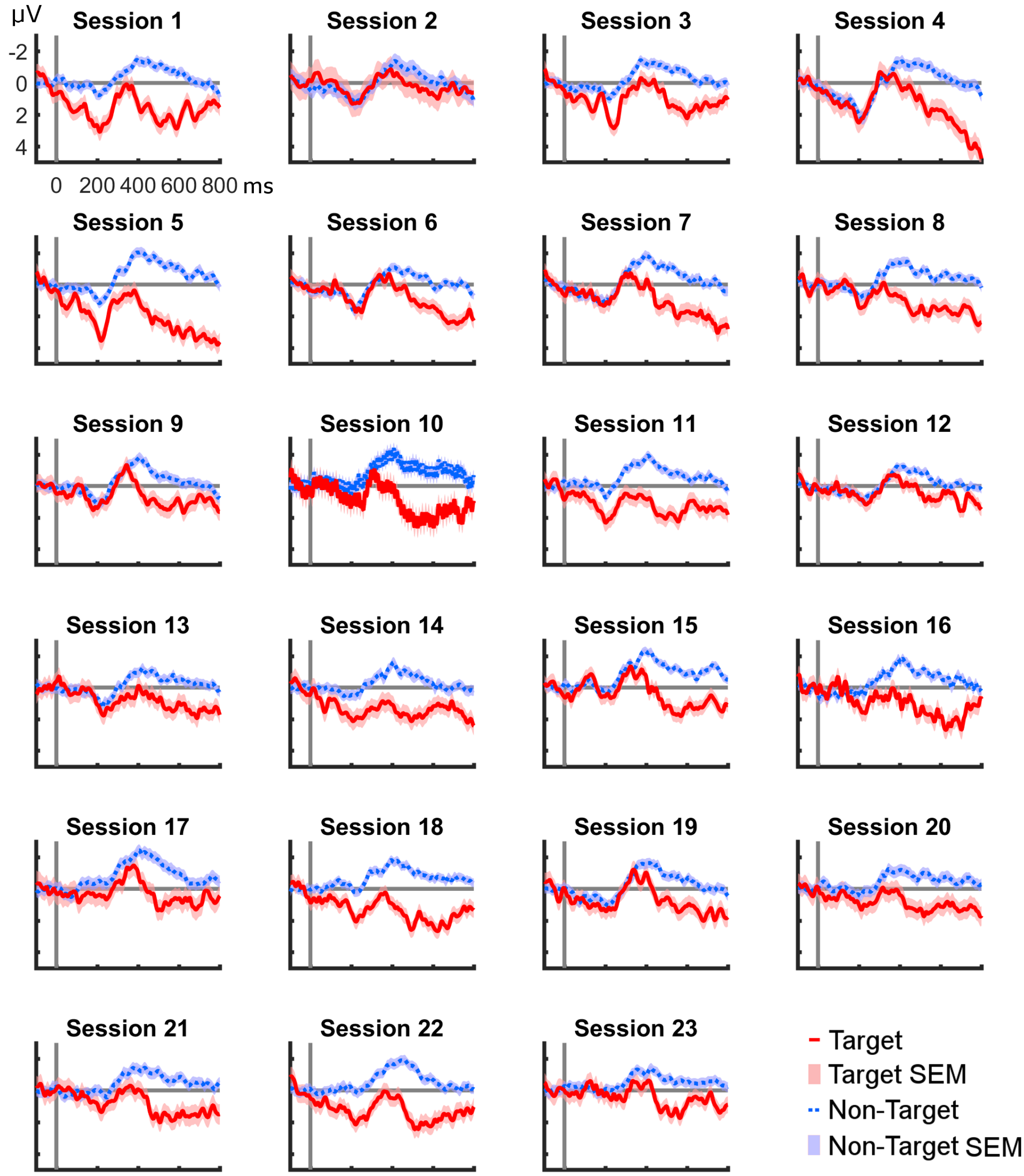

Figure A.7: Post-Stimulus EEG (Pz) from PT-N, by session.

Appendix B

Supplementary Tables

Table B.1: Study I: Individual BCI performances (all sessions).

	Online Accuracy [%]					Online ITR [bits/min]				
Participant	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5
1	70.5	96.7	59.1	75.0	75.0	2.65	4.67	2.27	4.79	3.83
2	88.9	80.5	77.8	93.8	83.3	5.33	3.38	5.34	6.31	5.25
4	93.8	96.7	91.2	93.5	96.7	6.31	10.50	11.53	9.39	21.00
5	63.6	86.8	96.7	89.5	93.8	2.89	3.72	8.40	8.15	12.61
6 ex	75.0	29.5	56.8	63.6	52.3	2.74	0.02	0.80	1.16	0.59
7 ex	45.5	38.6	29.5	36.4	43.2	0.34	0.17	0.02	0.11	0.39
8 ex	43.2	31.8	27.3	36.4	38.6	0.27	0.04	0.00	0.11	0.16
9	22.7	72.5	77.5	73.7	83.3	0.00	2.89	3.52	4.54	5.25
10 ex	47.7	38.6	25.0	45.5	34.1	0.42	0.16	0.00	0.34	0.09
11 ex	52.3	31.8	34.1	27.3	38.6	0.59	0.05	0.07	0.00	0.16
12	40.9	22.7	70.5	56.8	72.7	0.21	0.00	3.17	1.33	3.50
13	56.8	38.6	91.7	83.7	82.9	0.80	0.16	5.03	5.33	8.63
14	81.8	91.2	86.8	100.0	88.9	4.97	11.53	4.96	9.68	15.98
15	68.2	77.5	93.8	84.2	93.8	4.78	7.05	12.61	6.78	9.46
16 ex	68.2	66.7	77.5	65.9	59.1	1.44	2.23	3.02	2.58	1.51

Notes: ex = excluded. S = Session.

Table B.2: Study II: Individual BCI performances (all sessions).

	Online Accuracy [%]					Online ITR [bits/min]				
Participant	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5
1	87.5	95.8	91.7	100.0	95.8	3.81	5.09	4.40	6.04	5.09
2	83.3	58.3	70.8	83.3	100.0	3.28	1.09	2.02	3.28	6.04
3^{ex}	58.3	70.8	70.8	50.0	50.0	1.09	2.02	2.02	0.63	0.63
4	87.5	100.0	100.0	100.0	91.7	3.81	6.04	6.05	6.04	4.40
5	58.3	91.7	95.8	95.8	100.0	1.09	4.40	5.09	5.09	6.04
6	91.7	91.7	91.7	100.0	100.0	4.40	4.40	4.40	6.04	6.04
7	41.7	75.0	70.8	95.8	79.2	0.29	2.40	2.02	5.09	2.82
8	70.8	75.0	54.2	95.8	87.5	2.02	2.40	0.84	5.09	3.81
9	54.2	83.3	75.0	87.5	91.7	0.84	3.28	2.40	3.81	4.40
10	87.5	62.5	95.8	83.3	95.8	3.81	1.36	5.09	3.28	5.09
11	91.7	91.7	91.7	83.3	95.8	4.40	4.40	4.40	3.28	5.09
12	83.3	87.5	100.0	100.0	87.5	3.28	3.81	6.05	6.04	3.81
13	87.5	95.8	62.5	75.0	58.3	3.81	5.09	1.36	2.40	1.09
14	91.7	87.5	95.8	87.5	100.0	4.40	3.81	5.09	3.81	6.04
15	87.5	100.0	100.0	100.0	95.8	3.81	6.01	6.05	6.04	5.09
16	95.8	100.0	95.8	100.0	100.0	5.09	6.04	5.09	6.04	6.04
17	50.0	100.0	100.0	87.5	100.0	0.63	6.04	6.05	3.81	6.04
18	100.0	100.0	100.0	100.0	87.5	6.04	6.05	6.05	6.04	3.81
19	79.2	100.0	95.8	100.0	91.7	2.82	6.05	5.09	6.04	4.40
20	66.7	45.8	66.7	79.2	75.0	1.67	0.44	1.67	2.82	2.40

Notes: ^{ex} = excluded. S = Session.

Appendix C

Abbreviations

Abbreviation	Description
AT	Assistive technology
ALS	Amyotrophic lateral sclerosis
ALSFRS-R	Revised ALS functional rating score
BCI	Brain-computer interface
CLIS	Completely locked-in syndrome
EEG	Electroencephalography
ERP	Event-related potential
fMRI	Functional magnetic resonance imaging
ITR	Information transfer rate
LED	Light emitting diode
LIS	Locked-in syndrome
PEG	Percutaneous endoscopic gastrostomy
QUEST	Quebec User Evaluation of Satisfaction with assistive technology
SEP	Somatosensory evoked potentials
SMR	Sensory motor rhythms
swLDA	Stepwise linear discriminant analysis
TLX	Task load index
TUEBS	(Extended version of the QUEST 2.0)
UCD	User-centred design
UDP	User datagram protocol
VAS	Visual analogue scale

Declaration

Eidesstattliche Erklärungen nach §7 Abs. 2 Satz 3, 4, 5 der Promotionsordnung der Fakultät für Biologie

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation: **„Training Effects Of A Tactile Brain-Computer Interface System During Prolonged Use By Healthy And Motor-Impaired People"**, eigenständig, d. h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen, als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben. Die Regeln der Universität Würzburg über gute wissenschaftliche Praxis wurden eingehalten.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Weiterhin erkläre ich, dass bei allen Abbildungen und Texten bei denen die Verwertungsrechte (Copyright) nicht bei mir liegen, diese von den Rechtsinhabern eingeholt wurden und die Textstellen bzw. Abbildungen entsprechend den rechtlichen Vorgaben gekennzeichnet sind sowie bei Abbbildungen, die dem Internet entnommen wurden, der entsprechende Hypertextlink angegeben wurde.

Affidavit

I hereby declare that my thesis entitled: **„Training Effects Of A Tactile Brain-Computer Interface System During Prolonged Use By Healthy And Motor-Impaired People"** is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis. The work was performed under adherence to the regulations for good scientific practice by the University of Würzburg.

Furthermore I verify that the thesis has not been submitted as part of another examination process neither in identical nor in similar form.

Besides I declare that if I do not hold the copyright for figures and paragraphs, I obtained it from the rights holder and that paragraphs and figures have been marked according to law or for figures taken from the internet the hyperlink has been added accordingly.

Würzburg, den

Matthias Eidel