Brain-Computer Interfaces based on Event-Related Potentials:

toward fast, reliable and easy-to-use communication systems for people with neurodegenerative disease.

Gehirn-Computer Schnittstellen basierend auf Ereigniskorrelierten Potentialen:

Entwicklung von schnellen, zuverlässigen und leicht zu bedienenden Kommunikationssystemen für Menschen mit neurodegenerativer Erkrankung.

Doctoral thesis for a doctoral degree

at the Graduate School of Life Sciences,

Julius-Maximilians-Universität Würzburg,

Section *Neuroscience*

submitted by

Tobias Kaufmann

from

Rheinfelden, Germany

Würzburg, 2013

Submitted on: ....

Office stamp

Members of the *Promotionskomitee*:

Chairperson: Prof. Dr. Paul Pauli

Primary Supervisor: Prof. Dr. Andrea Kübler

Supervisor (Second): Prof. Dr. Wolfgang Rössler

Supervisor (Third): Prof. Dr. Klaus Schilling

Date of Public Defence: ....

Date of receipt of Certificates: ....

Dedicated to my parents

Gewidmet meinen Eltern

Danksagung (Acknowledgment)

Die hier vorliegende Dissertationsschrift habe ich am Lehrstuhl für Psychologie I der Universität Würzburg angefertigt. An dieser Stelle möchte ich mich bei all jenen bedanken, die mir im Laufe meiner Promotionszeit unterstützend zur Seite gestanden haben.

Ein herzliches Dankeschön an meine Supervisoren Prof. Dr. Andrea Kübler (Institut für Psychologie), Prof. Dr. Wolfgang Rössler (Biozentrum) und Prof. Dr. Klaus Schilling (Institut für Informatik) für die stets sehr freundliche Atmosphäre und die sehr ansprechenden Diskussionen unserer Meetings. Vielen Dank auch für die Begutachtung dieser Arbeit. Herrn Prof. Dr. Paul Pauli (Institut für Psychologie) danke ich für die Leitung meiner Promotionsverteidigung. Ein ganz besonderer Dank gilt Frau Prof. Dr. Kübler, die mich stets unterstützt und gefördert hat. Für die angenehme Arbeitsatmosphäre, die Unterstützung und die vielen Freiheiten, die sie mir als Chefin lies, bin ich sehr dankbar.

Zudem danke ich all meinen Kolleginnen und Kollegen am Lehrstuhl für Psychologie I, für das stets angenehme, häufig sehr lustige Arbeitsklima. Besonderer Dank an Dr. Sonja Kleih, die mir in vielen organisatorischen und projektbezogenen Belangen den Rücken frei hielt und auf deren freundschaftliche Unterstützung ich stets bauen konnte. Ebenfalls besonderer Dank an Dr. Stefan Sütterlin, der mir den Start in Würzburg sehr erleichtert hat und mir seit vielen Jahren sowohl als Wissenschaftler wie auch als guter Freund mit Rat und Tat zur Seite steht. Vielen Dank an Dr. Stefan Schulz, der meine Arbeit mit produktivem Input bereicherte, für die konstruktiven Diskussionen und die freundschaftliche Zusammenarbeit. Ich danke den Patientinnen und Patienten, die bereitwillig an meinen Studien teilgenommen haben, obschon dies für sie viel Aufwand bedeutete. Danke auch meinen DiplomandInnen und Praktikanten für das Interesse, das sie meiner Forschung entgegengebracht haben.

Auch gebührt Dank den Projektpartnern des EU Projektes TOBI (Tools for Brain Computer Interaction, FP7-224631), besonders den Kolleginnen und Kollegen der Technischen Universität Graz, die mich während einer Kollaborationsstudie sehr freundlich in ihre Gruppe aufgenommen haben. Meinen ganz besonderen Dank an Dr. Vera Kaiser und Dr. Günther Bauernfeind.

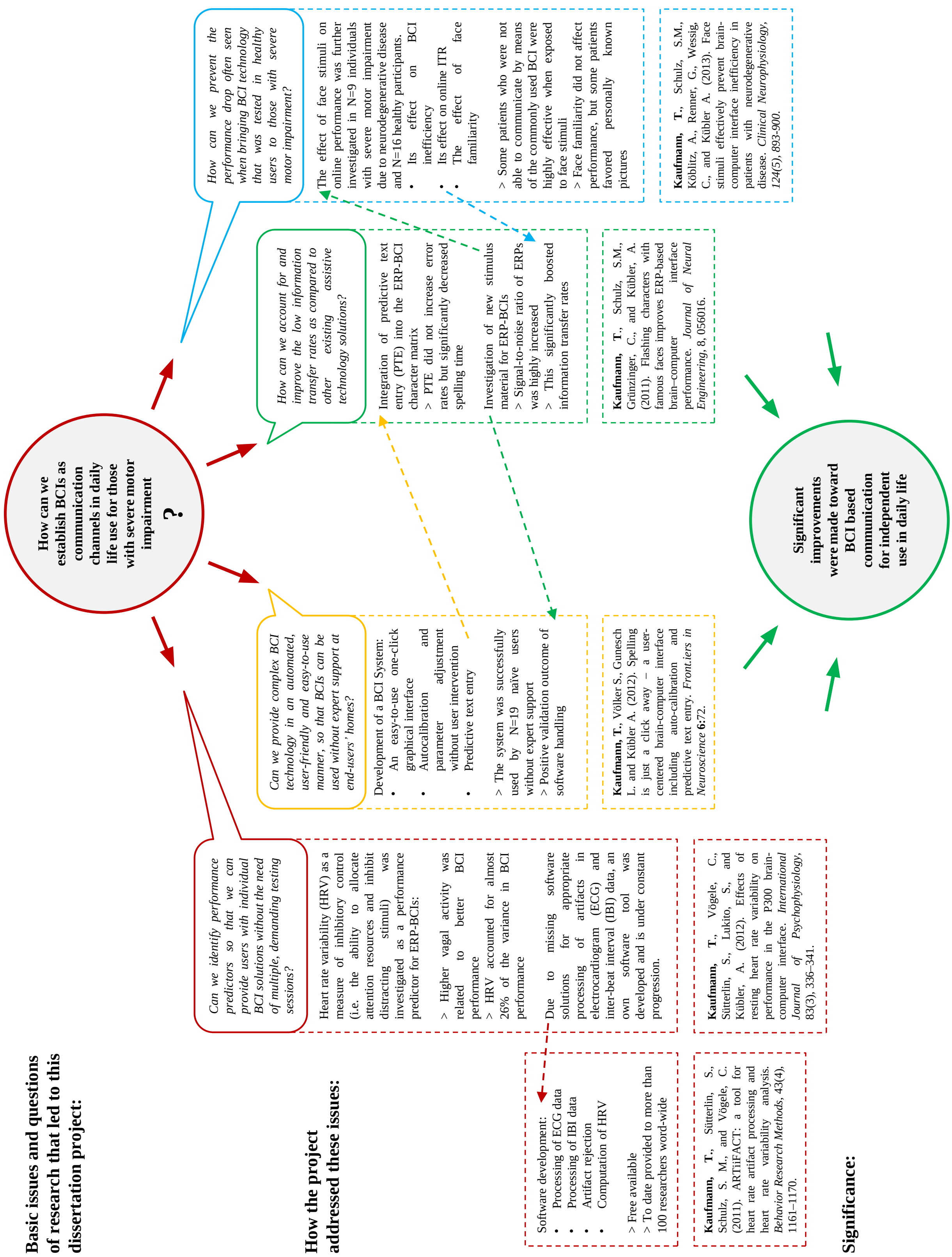
Danke auch dem Team der Graduate School of Life Science Würzburg, insbesondere Jennifer Braysher für ihr großes Engagement.

Neben der Unterstützung im Arbeitsumfeld habe ich auch von privater Seite viel Unterstützung erfahren. Ich möchte mich bei meinen Eltern, Geschwistern und Freunden ganz herzlich dafür bedanken. Ganz besonderer Dank gebührt meinen Eltern Waltraud und Herbert Kaufmann, die mich auf meinem Weg stets voll unterstützten und mir jegliche Freiheit ließen, diesen Weg selbst zu gestalten. Von ganzem Herzen danke ich auch meiner Freundin Pia März, die mich motiviert, mir immerzu viel Rückhalt gibt und die viel Verständnis für mein großes Arbeitspensum zeigte.

Vielen Dank an alle für die großartige Unterstützung!

Tobias Kaufmann

Graphical Abstract



Abstract

Objective: Brain Computer Interfaces (BCI) provide a muscle independent interaction channel making them particularly valuable for individuals with severe motor impairment. Thus, different BCI systems and applications have been proposed as assistive technology (AT) solutions for such patients. The most prominent system for communication utilizes event-related potentials (ERP) obtained from the electroencephalogram (EEG) to allow for communication on a character-by-character basis. Yet in their current state of technology, daily life use cases of such systems are rare. In addition to the high EEG preparation effort, one of the main reasons is the low information throughput compared to other existing AT solutions. Furthermore, when testing BCI systems in patients, a performance drop is usually observed compared to healthy users. Patients often display a low signal-to-noise ratio of the recorded EEG and detection of brain responses may be aggravated due to internally (e.g. spasm) or externally induced artifacts (e.g. from ventilation devices). Consequently, practical BCI systems need to cope with manifold inter-individual differences. Whilst these high demands lead to increasing complexity of the technology, daily life use of BCI systems requires straightforward setup including an easy-to-use graphical user interface that nonprofessionals can handle without expert support.

Research questions of this thesis: This dissertation project aimed at bringing forward BCI technology toward a possible integration into end-users' daily life. Four basic research questions were addressed: (1) Can we identify performance predictors so that we can provide users with individual BCI solutions without the need of multiple, demanding testing sessions? (2) Can we provide complex BCI technology in an automated, user-friendly and easy-to-use manner, so that BCIs can be used without expert support at end-users' homes? (3) How can we account for and improve the low information transfer rates as compared to other existing assistive technology solutions? (4) How can we prevent the performance drop often seen when bringing BCI technology that was tested in healthy users to those with severe motor impairment?

Results and discussion: (1) Heart rate variability (HRV) as an index of inhibitory control (i.e. the ability to allocate attention resources and inhibit distracting stimuli) was significantly related to ERP-BCI performance and accounted for almost 26% of variance. HRV is easy to assess from short heartbeat recordings and may thus serve as a performance predictor for ERP-BCIs. Due to missing software solutions for appropriate processing of artifacts in heartbeat data (electrocardiogram and inter-beat interval data), our own tool was developed that is available free of charge. To date, more than 100 researchers worldwide have requested the tool. Recently, a new version was developed and released together with a website (www.artiifact.de). (2) Furthermore, a study of this thesis demonstrated that BCI technology can be incorporated into easy-to-use software, including auto-calibration and predictive text entry. Naïve, healthy non-professionals were able to control the software without expert support and successfully spelled words using the auto-calibrated BCI. They reported that software handling was straightforward and that they would be able to explain the system to others. However, future research is required to study transfer of the results to patient samples. (3) The commonly used ERP-BCI paradigm was significantly improved. Instead of simply highlighting visually displayed characters as is usually done, pictures of famous faces were used as stimulus material. As a result, specific brain potentials involved in face recognition and face processing were elicited. The event-related EEG thus displayed an increased signal-to-noise ratio, which facilitated the detection of ERPs extremely well. Consequently, BCI performance was significantly increased. (4) The good results of this new face-flashing paradigm achieved with healthy participants transferred well to users with neurodegenerative disease. Using a face paradigm boosted information throughput. Importantly, two users who were highly inefficient with the commonly used paradigm displayed high accuracy when exposed to the face paradigm. The increased signal-to-noise ratio of the recorded EEG thus helped them to overcome their BCI inefficiency.

Significance: The presented work at hand (1) successfully identified a physiological predictor of ERP-BCI performance, (2) proved the technology ready to be operated by naïve nonprofessionals without expert support, (3) significantly improved the commonly used spelling paradigm and (4) thereby displayed a way to effectively prevent BCI inefficiency in patients with neurodegenerative disease. Additionally, missing software solutions for appropriate handling of artifacts in heartbeat data encouraged development of our own software tool that is available to the research community free of charge. In sum, this thesis significantly improved current BCI technology and enhanced our understanding of physiological correlates of BCI performance.

Zusammenfassung (German abstract)

Zielsetzung: Gehirn-Computer Schnittstellen (engl. Brain-Computer Interface, BCI) bilden einen muskel-unabhängigen Interaktionskanal, was sie besonders für Menschen mit starken, motorischen Einschränkungen wertvoll macht. Daher wurden verschiedene BCI-Systeme und -Anwendungen als unterstützende Technologien (UT) für diese Patienten vorgeschlagen. Das am häufigsten verwendete System zu Kommunikationszwecken basiert auf ereigniskorrelierten Potentialen (EKP) des Elektroenzephalogramms (EEG) und ermöglicht es Buchstabe für Buchstabe zu kommunizieren. Zum derzeitigen Stand der Technik sind Berichte über alltägliche Verwendung von BCI-Systemen jedoch selten. Zusätzlich zu dem hohen Präparationsaufwand, der mit dem Messen eines EEGs verbunden ist, ist eine der Hauptursachen der verhältnismäßig geringe Informationstransfer im Vergleich zu anderen existierenden UT-Lösungen. Zudem wird häufig beobachtet, dass das BCI-Kontrollvermögen bei Patienten deutlich niedriger ist als bei gesunden Nutzern. Das EEG von Patienten weißt häufig ein niedrigeres Signal-zu-Rausch-Verhältnis auf und die Erkennung von Hirnantworten kann durch interne (z. B. Spasmus) oder externe Artefakte (z. B. von einem Beatmungsgerät herrührend) zusätzlich erschwert werden. Somit müssen praxisrelevante BCI-Systeme mit einer Vielfalt von interindividuellen Unterschieden klar kommen. Obschon diese hohen Anforderungen zu einer zunehmend komplexeren Technologie führen, erfordert der Alltagsgebrauch von BCI-Systemen einen einfachen Aufbau inklusive leicht zu bedienender, grafischer Benutzeroberfläche, die von Laien, ohne Unterstützung von Experten, bedient werden kann.

Forschungsfragestellungen dieser Dissertationsschrift: Das Dissertationsprojekt bezweckte die Weiterentwicklung von BCI-Systemen, um den Weg für eine Integration in das tägliche Leben von Benutzern zu bahnen. Vier grundlegende Forschungsfragestellungen wurden adressiert: (1) Können Prädiktoren des BCI-Kontrollvermögens gefunden werden, sodass Nutzer mit individuell angepassten BCI-Lösungen versorgt werden können, ohne dass mehrfache, anstrengende Testsitzungen notwendig sind? (2) Kann komplexe BCI-Technologie auf automatisierte, nutzerfreundliche und leicht zu bedienende Weise bereitgestellt werden, sodass BCIs, ohne die Unterstützung von Experten, zuhause verwendet werden können? (3) Wie kann man den im Vergleich zu anderen existierenden UT-Lösungen niedrigen Informationstransferraten begeg-

nen und sie erhöhen? (4) Wie kann ein Abfall des Kontrollvermögens verhindert werden, der häufig in Erscheinung tritt, wenn man BCI-Technologie, die an Gesunden getestet wurde, zu Patienten mit starken motorischen Einschränkungen bringt?

Ergebnisse und Diskussion: (1) Herzratenvariabilität als Index des inhibitorischen Kontrollvermögens (die Fähigkeit Aufmerksamkeitsressourcen bereitzustellen und ablenkende Reize zu inhibieren) wurde zu EKP-BCI-Kontrollvermögen in einen signifikanten Bezug gesetzt und klärte beinahe 26% der Varianz auf. HRV ist leicht aus kurzen Aufzeichnungen des Herzschlags zu erheben und könnte daher als Prädiktor des EKP-BCI-Kontrollvermögens dienen. Aufgrund fehlender Softwarelösungen für angemessene Artefaktbehandlung in Aufzeichnungen des Herzschlags (Elektrokardiogramm und Interbeat-Intervall Daten) wurde ein eigenes Programm entwickelt, das frei erhältlich ist. Bis heute wurde es von mehr als 100 Wissenschaftlern weltweit angefordert. Unlängst wurde zudem eine neue Version entwickelt, die zusammen mit einer Website veröffentlicht wurde (www.artiifact.de). (2) Des Weiteren zeigte eine Studie dieser Dissertation, dass BCI-Technologie samt Auto-Kalibration und Wortvervollständigung in eine leicht zu bedienende Software integriert werden kann. Ungeschulte, gesunde Probanden waren in der Lage die Software ohne Unterstützung von Experten zu bedienen und buchstabierten mit dem auto-kalibrierten BCI erfolgreich Wörter. Sie gaben an, dass die Bedienung der Software leicht zu tätigen sei und dass sie in der Lage wären, das System anderen zu erklären. Jedoch muss zukünftige Forschung klären, ob sich die Ergebnisse auf Patienten übertragen lassen. (3) Das häufig verwendete EKP-BCI-Paradigma wurde signifikant verbessert. Statt - wie normalerweise getätigt - visuell präsentierte Buchstaben einfach aufleuchten zu lassen, wurden Bilder berühmter Gesichter als Stimulationsmaterial verwendet. Hierdurch wurden spezifische Gehirnpotentiale der Gesichtserkennung und -verarbeitung ausgelöst. Das ereigniskorrelierte EEG wies daher ein höheres Signal-zu-Rausch-Verhältnis auf, was die Detektion von EKPs stark vereinfachte. Infolgedessen war das BCI-Kontrollvermögen signifikant erhöht. (4) Die guten Ergebnisse dieses neuen Gesichter-Stimulus-Paradigmas, die mit gesunden Probanden erreicht wurden, ließen sich gut auf Patienten mit neurodegenerativen Erkrankungen übertragen. Die Verwendung eines Gesichter-Paradigmas erhöhte den Informationstransfer erheblich. Zwei Nutzer, die sehr ineffizient im Umgang mit dem herkömmlichen BCI-System waren, erreichten zudem ein hohes Kontrollvermögen mit dem Gesichter-Paradigma. Das erhöhte

Signal-zu-Rausch-Verhältnis des aufgezeichneten EEGs half ihnen somit ihre BCI Ineffizienz zu überwinden.

Signifikanz: Die hier vorgestellte Arbeit (1) identifizierte einen physiologischen Prädiktor des EKP-BCI-Kontrollvermögens, (2) zeigte, dass die Technologie bereit ist für die Verwendung durch ungeschulte Laien ohne Unterstützung von Experten, (3) verbesserte das herkömmliche Kommunikationsparadigma signifikant, und (4) zeigte hierdurch einen Weg auf, die BCI Ineffizienz von Patienten mit neurodegenerativen Erkrankungen effektiv zu verhindern. Die fehlende Softwarelösung zur angemessenen Behandlung von Artefakten in Aufzeichnungen des Herzschlags animierte zudem zur Entwicklung einer eigenen Anwendung, die der Wissenschaftsgemeinschaft kostenfrei zur Verfügung gestellt wird. Zusammenfassend kann gesagt werden, dass dieses Dissertationsprojekt somit derzeitige BCI-Technologie signifikant verbesserte und unser Verständnis physiologischer Korrelate des BCI-Kontrollvermögens erweiterte.

Contents

List of Figures

List of Abbreviations

Abbreviation	Description
ALS	Amyotrophic lateral sclerosis
	fALS: familial ALS
	sALS: sporadic ALS
AT	Assistive technology
BCI	Brain-computer interface
BMI	Body mass index
	In the literature BMI also abbreviates brain-machine interfaces
BOLD	Blood oxygen-level dependent
CLIS	Complete locked-in syndrome
ECG	Electrocardiography
ECoG	Electrocorticography
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
ERP	Event-related potential
ERD	Event-related desynchronization
ERS	Event-related synchronization
fMRI	Functional magnetic resonance imaging
HRV	Heart rate variability
Hb	Hemoglobin
	Oxy-Hb: oxygenated hemoglobin
	Deoxy-Hb: deoxygenated hemoglobin
IBI	Inter-beat interval
ISI	Inter-stimulus interval
IQR	Inter-quartile range

ITR	Information transfer rate
LIS	Locked-in syndrome
LF/HF	Low frequency / high frequency
LFP	Local field potentials
MI	Motor imagery
MED	Maximum expected difference
	i.e. maximum expected beat difference of valid beats
MAD	Minimum artificial difference
	i.e. minimum expected difference of artificial beats
NIRS	Near infrared spectroscopy
NoS	Number of sequences
OCI	*Optimized Communication Interface*
PEG	Percutaneous endoscopic-controlled gastrostomy
PTE	Predictive text entry
RSA	Respiratory sinus arrhythmia
SCP	Slow cortical potentials
SMA	Spinal muscular atrophy
SMR	Sensory motor rhythms
SNR	Signal-to-noise ratio
SSVEP	Steady-state visual-evoked potentials
SSSEP	Steady-state somatosensory-evoked potentials
SSAEP	Steady-state auditory-evoked potentials
SWLDA	Stepwise linear discriminant analysis

1 Introduction

In November 2012, I received an email from a client for who we tried to develop a brain-computer interface specifically adjusted to his needs. He is blind, severely motor-impaired and has no means of verbal communication but since few years, he is again able to control a computer. He started documenting the history and progression of his disease and sent me some of his notes.

After an accident 21 years ago, he had been in the locked-in state, not able to control any muscle apart from his blind eyes voluntarily. Communication would have been possible by asking closed questions so that he could answer in binary codes (yes=eyelift, no=eye down). Unfortunately, neither medical doctors nor caregivers thought of this possibility and so he was bound to silence although he was fully conscious and himself aware of this communication possibility. His short statements give a moving insight into a world we can hardly imagine. Not only had he to process the burden of his acquired disabilities mentally, but also had he to endure the cruelty of caregivers.

Two months later, a caregiver finally had the idea of communicating by reading eye movement. Although this brought back a means of communication, still he could communicate only if someone took the time to talk to him. In the process of his physical rehabilitation, he gained back some muscular control. Today he is able to generate a sound if he wants to communicate and to use his tongue and his right arm to communicate commands in a partner scanning approach.

Reading his lines reminds us that communication is a privilege that must not be taken from anyone and that no effort in giving back a means of independent communication should be spared.

(With permission from the client to report on these incidences.)

In the following, an overview is provided of some typical diseases and damages to neuromuscular pathways that may lead to severe motor impairment and thus entail loss of verbal communication ability (chapter 1.1). Thereafter, brain-computer interfaces are introduced with particular focus on their use as an assistive communication technology (chapters 1.2-1.3). Finally, basic issues leading to the work at hand are pointed out, connected to the model of BCI control (introduced in chapter 1.4) and a brief compendium is provided of how these issues were addressed in this thesis (chapter 1.5).

1.1 Neurodegenerative diseases and brain damages that may lead to a locked-in syndrome

In the following chapter a short overview is given of amyotrophic lateral sclerosis (ALS), spinal muscular atrophy (SMA) and brainstem stroke, as patients with these diseases and damages are examples of people who may in the future benefit from direct brain-computer based communication technology. In particular patients that can be classified under the term ”Locked-in syndrome” (LIS; Plum and Posner 1966) may benefit from such technology. In general, it describes a state of quadriplegia and paralysis of the lower cranial nerves except for those involved in control of vertical eye movements which retain functionality. For more accurate classification, Bauer, Gerstenbach and Rumpl (1979) broke down the general definition of LIS into three distinct stages based on the level of motor impairment, being

1. *Classical LIS*: Total paralysis except for vertical eye movements and blinking.
2. *Incomplete LIS*: Severe paralysis but retaining muscle control in addition to vertical eye movement and blinking.
3. *Total LIS*: Total paralysis including occular muscles.

Importantly, patients in all stages of LIS retain consciousness (Bauer et al., 1979). Most often LIS results from vascular pathology or traumatic brain injury but may also result from neurodegenerative diseases such as ALS or SMA (for review, Laureys et al., 2005; Bruno et al., 2009). To account for different causes of LIS (and thus different progressions of disease), Kübler and Birbaumer (2008) suggested distinguishing between two stages, being

1. *LIS*: Almost complete paralysis except for some preserved muscle control.
2. *CLIS*: Complete locked-in syndrome, i.e. complete paralysis.

1.1.1 Amyotrophic lateral sclerosis (ALS)

Amyotophic lateral sclerosis (ALS), the most common neurodegenerative disease of the motor system, was first described by Charcot and Joffroy (1869). It is caused by a progressive degeneration of higher motor neurons in the motor cortex and of lower motor neurons in the brainstem and spinal cord (for review, Kiernan et al., 2011; see figure 1). However, recent genetic research points to a coupling of ALS to frontotemporal degeneration, thereby indicating that other cell types are involved as well (for review, Ludolph et al., 2012). These findings are further supported by imaging studies displaying affection of frontal lobes in both, demented and non-demented ALS patients (e.g., Abrahams et al., 1995; Kiernan and Hudson, 1994; Kato et al., 1993). Cognitive dysfunction was found to appear in the early stage of ALS and particularly in the bulbar form (Schreiber et al., 2005). However, the dominant progression appears to be motor degeneration, which was found to progress faster than frontal deficits (Schreiber et al., 2005).

A meta-analysis by Byrne and colleagues (2011) estimated familial ALS (fALS; inherited form) to 5.1% of cases, i.e. in most cases ALS is sporadic (sALS) and not related to familial anamneses. Usually, fALS and sALS are clinically indistinguishable (for review, Hand and Rouleau, 2002). The mean age of disease onset is slightly lower for fALS than for sALS (for review, Hudson, 1981). Disease duration (i.e. from onset of recognized motor deficits to death) was related to onset age but not to gender (Eisen et al., 1993). Around 50% of ALS patients die within 30-36 months from the first symptoms (for review, Talbot, 2009; Mitchell and Borasio, 2007). A pooled analysis of registered ALS cases across European countries revealed an incidence rate of 2.16 cases per 100.000 persons per year (Logroscino et al., 2010) and compares to estimations for the US (McGuire et al., 1996). The male to female incidence ratio was calculated to be 1.4:1, i.e. men are more often affected than women (Logroscino et al., 2010).

ALS is clinically classified in terms of four subtypes depending on the primarily affected regions (for review, Mitchell and Borasio, 2007):

1. *Bulbar*: Patients present impaired speech (dysarthria) or difficulties in chewing and swallowing (dysphagia) attributed to lower motor neurons degeneration (bulbar palsy) or upper motor neurons degeneration (pseudobulbar palsy), or both.
2. *Cervical*: Patients present impairment in upper-limb extremities due to lower, upper or both motor neurons degeneration.
3. *Lumbar*: Patients present impairment in lower-limb extremities due to lower motor neurons degeneration.
4. *Thoracic*: Patients present impairment in respiration.

Bulbar-onset ALS occurs in about 25% of cases, cervical- and lumbar-onset ALS in about 70%, whereas the thoracic subtype occurs very rarely only in about 5% of cases (Kiernan et al., 2011).

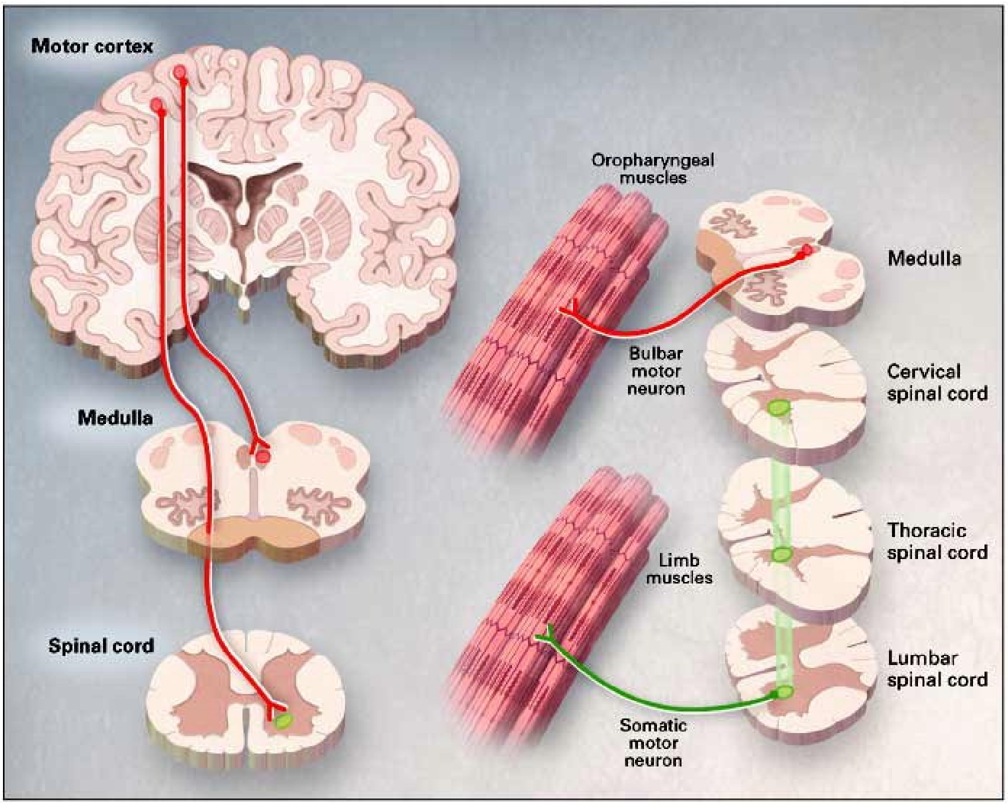


Figure 1: Causes of amyotrophic lateral sclerosis. Degeneration of motor neurons in the cortex, brainstem and spinal cord affect muscle control. Figure reproduced with permission from Rowland and Shneider (2001), Copyright Massachusetts Medical Society. The publication is available online at http://www.nejm.org.

ALS subtypes differ in terms of primarily affected regions but share many symptoms that may occur in different order and characteristics varying from case to case. Typical symptoms are muscular atrophy, muscle fasciculation, cramps or even spasm, dysarthria, dysphagia, dyspnea (for review, Borasio and Voltz, 1997). ALS patients were found to have a chronically deficient energy intake that is independent of dysphagia and may be related to an increased respiratory effort or demands put on the remaining, functional muscles (Kasarskis et al., 1996). Thus, adequate nutrition can be guaranteed through percutaneous endoscopic-controlled gastrostomy (PEG; Mazzini et al., 1995) which also prevents risk of pulmonary aspiration (Kasarskis et al., 1996). Mazzini and colleagues (1995) effectively demonstrated in a group of N=31 ALS patients that PEG could improve their body mass index (BMI) as compared to a significant decrease of BMI in a control group (N=35). Importantly, survival was prolonged in the PEG group.

Most patients with ALS die due to respiratory failure (Hardiman et al., 2011). With the onset of respiration problems, patients are artificially ventilated utilizing non-invasive oxygen-support, which was found to prolong life of patients with normal to moderately impaired bulbar functions and to relieve symptoms (Radunovic et al., 2009). However, with progression of disease, this treatment loses feasibility and patients are required to decide if invasive, mechanical ventilation is requested for prolonging their life (Hardiman et al., 2011). In case patients decide against invasive ventilation, palliative care can be utilized for relief of symptoms in the last stage of the disease (Borasio and Voltz, 1997).

1.1.2 Spinal muscular atrophy (SMA)

Spinal muscular atrophy (SMA) is a genetic neurodegenerative disease affecting the lower motor neurons thereby leading to severe paralysis (for review e.g., Talbot and Davies, 2001; Cifuentes-Diaz et al., 2002). According to International SMA Consortium (Munsat, 1991, cf., e.g., Cifuentes-Diaz et al., 2002) the autosomal-recessive SMA is classified into three most common clinical subgroups (type I-III). A very rare, forth subgroup (type IV) is often distinguished for adult onset SMA.

1. *Type I*: Also referred to as Werding-Hofmann disease. Onset within the first 6 months. Patients can never sit unsupported.

2. *Type II*: Also referred to as intermediate SMA. Onset between 6 to 18 months after birth. Patients are able to sit unsupported but never able to stand or walk.

3. *Type III*: Also referred to as Kugelberg-Welander or Juvenile SMA. Onset usually after 18 month and within the first three years of life. Patients are able to stand and walk alone but perceive muscle weakness, which often entails difficulties in motor actions and falling accidents. Depending on the age of onset, type IIIa (before 3 years) and IIIb (after 3 years) are distinguished.

4. *Type IV*: Also referred to as adult onset SMA. Onset typically >30 years.

In a study with N=445 SMA patients, Zerres and Rudnik-Schöneborn (1995) reported survival properties at 2, 4, 10 and 20 years of age. Only 18% of SMA type I patients survived four years and none reached an age of 20. Patients with type II showed highly increased survival rates with 98% surviving 10 years and 77% reaching 20 years. As life duration is typically merely affected in type III patients the authors assessed ambulatory abilities for these patients. Type IIIb patients remained ambulatory for a longer time with 67% of cases after 40 years as compared to 34% of type IIIa patients. Prevalence of SMA has been estimated to 1.2 cases in 100.000 inhabitants with an incidence rate of 1 in 24.100 live births (Pearn, 1978).

1.1.3 Brainstem stroke

Impairment in the blood support of the vertebral arteries and its branches can lead to strokes in the mid-brain, the pons and the medulla oblongata and may occur due to embolism, branch or small artery disease or stenosis (for further reading, e.g. Grehl and Reinhardt, 2008). From a data base analysis with N=1000 stroke patients, Bogousslavsky, van Melle and Regli (1988) reported 48% of vertebrosilar infarkts being in the brainstem with 27% mainly in the pons, 14% in the medulla and 7% in the midbrain. In a study with N=407 brainstem stroke patients, embolism was found to be the most common reason (estimated to 40-54% of cases; with 24-33% cardiac sources, 14-18% intraarterial sources, 2-3% both; Caplan et al., 2004). Stenosis of large artery leading to hemodynamic ischemia was the possible reason in 32-35% of cases and branch artery disease in 14-17%. Glass and colleagues (2002) report a low mortality rate of 3.6% 30 days after the stroke. Major disabilities were found in 18% of patients, whereas 28% showed no disability. The majority presented with minor disabilities (51%). In line with Caplan and colleagues (2004), the stroke was caused by embolism in 55% of those patients with poor outcome, i.e. embolism is associated with highest risk of poor outcome. As already stated in chapter 1.1, brainstem stroke is the most common cause of locked-in syndrome (for review, e.g., Smith and Delargy, 2005; Laureys et al., 2005).

1.2 A prospect on direct brain-computer interfacing (BCI)

Thanks to medical progress, good survival rates for locked-in patients are achieved, however, establishing a reliable communication channel for them appears to remain one of the major issues to deal with (Dollfus et al., 1990). Usually, communication is muscle-dependent, i.e. assistive communication devices or communication partners need to rely on whatever muscular control is individually possible.

Locked-in patients usually retain vertical eye movements that allow for binary (yes/no) communication (e.g., Plum and Posner, 1982, pg. 9). In a so-called partner-scanning approach, a communication partner asks questions in the closed format, so that patients can respond with yes (e.g. lift of eye lid or looking up) or no (e.g. eye lid down or eyes remain in center position). Based on this technique, a basically unlimited communication can be developed as well. For example in the approach depicted in figure 2, letters of the alphabet are grouped into four categories based on their frequency of occurrence in daily language. The communication partner first reads out the categories and the patient selects one category (e.g. by eyelift). Subsequently, each letter of this category is read out to the patient thereby allowing for selection of a particular letter. The communication partner notes down this letter and repeats the same procedure for the next character until the patient communicated the intended words. Fortunately, this approach can be fastened once the communication partner is able to guess the intended word or even the whole statement. Laureys and colleagues (2005) provide an overview of diverse, similar approaches.

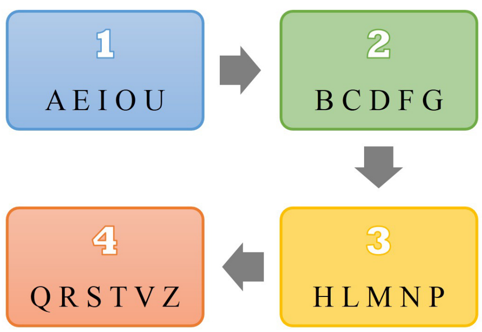


Figure 2: Exemplary illustration of the communication solution used by an Italian locked-in patient visited in the period of this dissertation project. Four categories comprise letters ranked by their frequency of occurrence in daily Italian language use. The patient first selects the category and subsequently the intended character. A communication partner notes down character by character until the patient's statement is complete.

Partner scanning unfortunately does not provide a means of independent communication, i.e. communication is only possible if someone reads a patient's eye movements. Assistive technology devices can account for this issue. For example, when connecting a blink detection sensor to a smart device that reads out characters like a real communication partner and processes input from the sensor, patients can communicate independently. Depending on the retaining muscle control other input sensors may be feasible such as eye-tracking (detection of vertical and horizontal pupil movements), button press (e.g. with tiny finger movement), tongue joystick, head or chin mouse. It is possible to integrate a realistic voice synthesizer into the system so that patients regain a voice that is personally associated with them (e.g., Doble et al., 2003). Furthermore, such devices are not restricted to basic communication but may also allow for computer and environmental control, e.g. control of a smart home, web browser or email application. Independence gained by using assistive technology devices is one factor positively influencing patients' perceived quality of life (e.g. Abresch et al., 1998; Kübler et al., 2005b), which is in particular affected by retaining communication abilities (e.g. Bach, 1993; Lulé et al., 2008). However, muscle-dependent communication entails some major difficulties to deal with.

1. *Fatigue*: Frequent use of the muscle with retained control may be fatiguing.
2. *Reliability*: Reliability of muscles may be limited or deteriorate with progression of disease.
3. *Speed*: If restricted to binary communication, speed is usually slow.
4. *Interference*: The muscle may be needed for other tasks, e.g. blink reflex in the case of eye-controlled devices, leading to miss-detection of commands.

Thus, researchers have been heavily investigating the direct use of brain signals for establishing muscle-independent communication and control. The term brain-computer interface (BCI) has been introduced to describe an approach in which brain signals are recorded, processed and directly translated into a device command without use of muscle control (Vidal, 1973). BCIs have already been proven as a functional means of communication in all stages of LIS except for CLIS (for review, e.g., Kübler and Birbaumer, 2008; Birbaumer et al., 2008). It is a big ambition of this field to establish communication with patients in the latter state. However, already for those patients retaining muscle control, BCIs may comprise a beneficial alternative to muscle-dependent devices. Current research is thus investigating possibilities to make BCIs faster, more reliable and thereby practical for daily life home use.

The following section will give an overview on current BCI technology, particularly focusing on non-invasive BCI-based communication devices.

1.3 Brain-Computer Interface (BCI)

1.3.1 Brain signals and recording techniques

Brain-computer interfaces (BCI) provide a direct connection between the brain and technical devices by means of signals recorded from inside (invasive) or outside the brain (non-invasive); for review see e.g., Kübler et al., 2001a; Birbaumer and Cohen, 2007; Kübler and Müller, 2007; Birbaumer et al., 2008; Allison et al., 2012a; Wolpaw and Wolpaw, 2012a). Signals are acquired, processed (digitized, amplified, filtered, features extracted and classified) and translated into a command for controlling devices such as computer applications or prosthetic devices. Outcome of classification is fed back to the user thereby completing the brain-computer interface circle (figure 3).

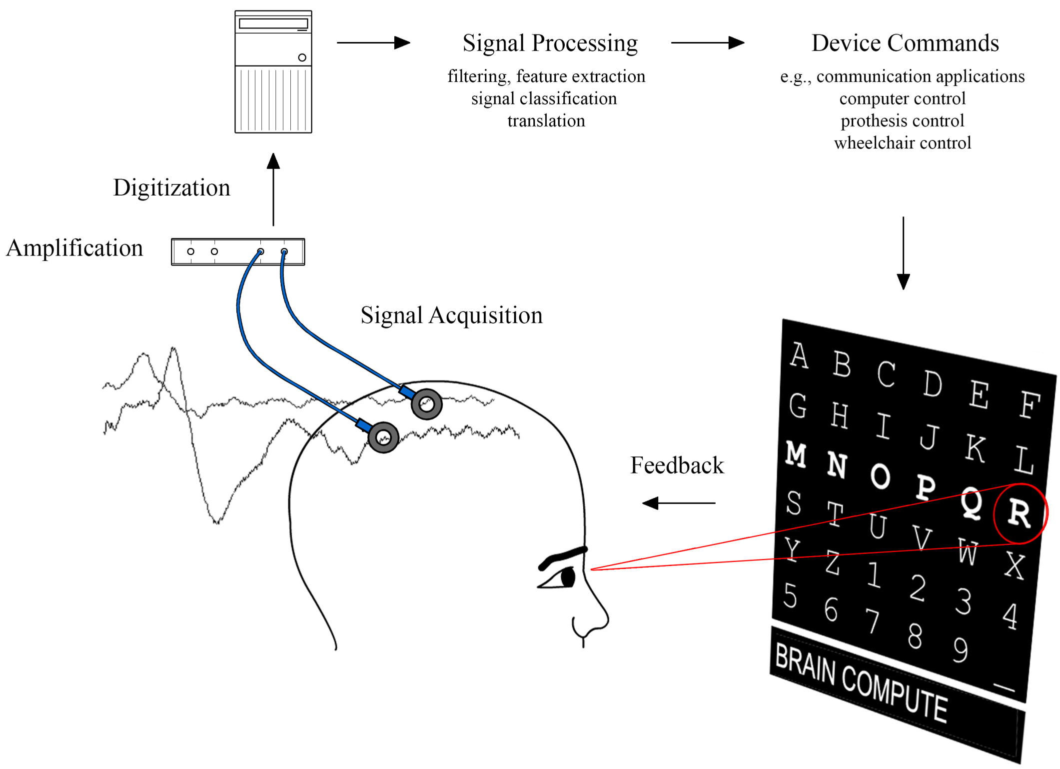


Figure 3: The brain-computer interface circle. Signals are acquired, processed and translated into a command. Feedback is provided to the user thereby closing the loop.

Recording signals from the human brain means making a compromise in terms of (1) good spatial resolution, (2) good temporal resolution and (3) extent to which the brain can be imaged (for review, e.g. Ramsey, 2012). Different methods have been proposed that greatly differ in these terms and the best method is to be selected based on the requirements of the planned investigations (Wolpaw and Wolpaw, 2012b).

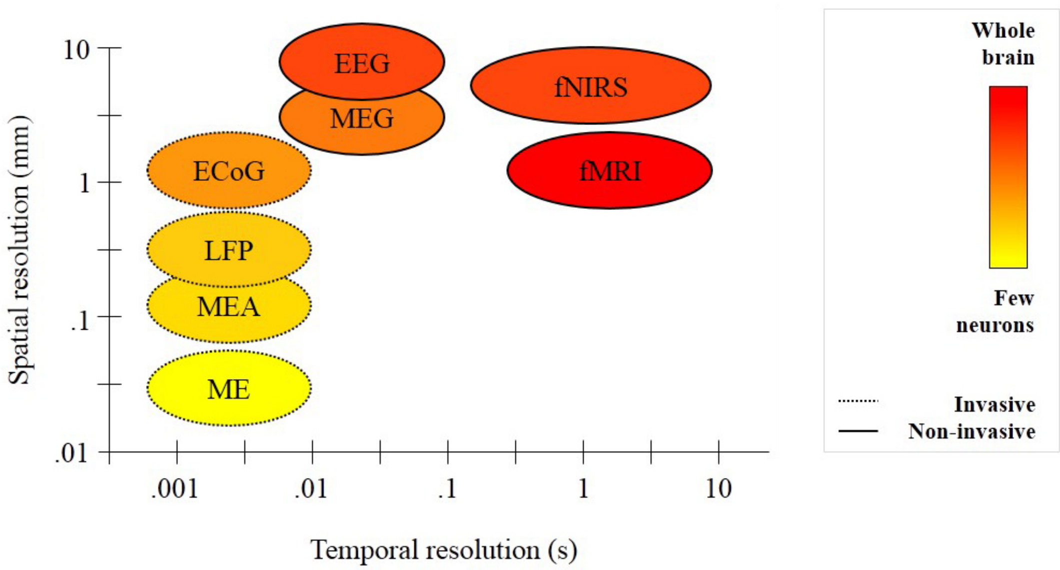


Figure 4: Comparison of brain signal recording techniques in terms of (1) spatial resolution, (2) temporal resolution and (3) extent to which the brain can be imaged. Please note that the values are an approximation and not based on accurate values. They may well differ from system to system. Thus, this figure is an illustration for approximate comparison only. The values for spatial and temporal resolution were taken from (van Gerven et al., 2009)

.

Invasive BCIs acquire signals from single neurons (single unit recording of spike trains), from small sets of neurons (local field potentials, LFP) or from electrode arrays placed on the cortical surface (electrocorticography, ECoG) (for review, e.g., Otto et al., 2012). As signals can be obtained closer to the source as compared to non-invasive recording on top of the head, invasive BCIs usually provide better signal quality and higher spatial resolution. One of the main challenges, however, is to obtain good signal quality over years as invasive electrodes often bear risk of short durability and bacterial infections (Otto et al., 2012). As this work solely focused on non-invasive BCIs, invasive BCIs are not introduced in depth. For comparison of invasive and non-invasive prospects, see Birbaumer (2006).

Non-invasive BCIs constitute an alternative of lower risk at the cost of lower signal resolution. As multiple degrees of freedom can be obtained even from non-invasive recordings, it remains to be shown if invasive technology can provide better results in terms of BCI accuracy, reliability and practical use (Wolpaw and Wolpaw, 2012b; Donoghue, 2012). For communication purpose non-invasive BCIs currently are the method of choice (Birbaumer, 2006). Two types of signal are distinguished:

1. *Hemodynamic brain response*: Hemodynamic recording techniques measure the blood-oxygen-level dependent (BOLD) response in brain regions under investigation (Ogawa et al., 1990). Approximately one second after onset of an increasing neural activity, oxygen consumption from surrounding capillaries increases. Consequently, the capillaries widen, thereby triggering a flow of oxygenated hemoglobin toward the source of increased neural activity. Each hemoglobin carries four O_2 molecules (oxy-Hb) and returns to the deoxygenated state (deoxy-Hb) once O_2 is delivered. Importantly, these two forms of hemoglobin show differences in color (i.e. light absorption; measured with functional near infrared spectroscopy, NIRS) and magnetism (measured with functional magnet resonance imaging, fMRI). These properties are utilized to measure the concentration ratio of oxy and deoxy-Hb, which directly reflects the level of neural activity. In sum, hemodynamic responses reflect neural activity but are of low temporal resolution due to the time delay of the BOLD response. For review on hemodynamic brain response, recording techniques and its use in BCI, see e.g., Ramsey, 2012; Sitaram et al., 2012.

2. *Electric brain response*: Unlike hemodynamic responses, electrical brain recordings directly access activity of neural assemblies. Thus, it is possible to record activity in real-time. An electroencephalogram (EEG) is obtained by recording voltage differences between electrodes at position of interest from the top of the head and a reference electrode usually placed on the nose, mastoids or ear-lobes (one further electrode is required to build a physical ground). The signal-to-noise ratio of EEG recordings is negatively affected by noise of tissues between the signal source and the electrodes. For review on electric brain response, recording techniques and its use in BCI, see e.g., Pizzagalli, 2007; Srinivasan, 2012.

Due to its high temporal resolution, low risk and convenient portability of the equipment, EEG was the recording technique of choice in this thesis. Furthermore, the straightforward setup renders EEG of practical value for a potential daily life BCI application.

In the following, I will give a short overview of the potential of some EEG-based BCI approaches to draw the reader's attention to those solutions that were not in the scope of this thesis, thereby allowing comparison of the presented results with other possible approaches. Subsequently, a separate section is dedicated to event-related potential-based BCIs, which were under investigation in this thesis (section 1.3.2).

1. *Slow cortical potentials (SCP)*: SCPs are slow potential shifts in the EEG which occur in response to imagination (e.g., movement or emotional situations) or cognitive tasks (e.g., waiting for a go cue; for review, e.g., Birbaumer et al., 1990). In an operant conditioning paradigm, users are trained to actively modulate their SCP amplitude while perceiving feedback on classification outcome. SCP amplitude can either be modulated downwards or upwards compared to a baseline condition that is estimated at the beginning of each trial. SCP based communication has been demonstrated in severely disabled locked-in patients (Kübler et al., 1999; Birbaumer et al., 1999). Unfortunately, SCP-BCIs are rather slow and increasing speed was found to be limited, i.e. very short intervals of SCP modulation were reported to be exhausting (Kübler et al., 1999). Furthermore, intensive training is required whilst accuracies were moderate (around 70-85% control, Kübler et al., 1999; Hinterberger et al., 2004).

2. *Sensory motor rhythms (SMR)*: When analyzing the frequency spectrum of the EEG recorded over sensory-motor areas during resting state of participants, it can be observed that rhythms in the mu (9-12Hz) and beta (20-30Hz) range are highly synchronized (Pfurtscheller and Neuper, 1997; Pfurtscheller et al., 1997; for a recent review Pfurtscheller and McFarland, 2012). Due to their origin in sensory-motor areas they are called sensory-motor rhythms (SMR). When participants perform a motor action, these rhythms desynchronize (figure 5). As this pattern of synchronization and desynchronization can be regulated voluntarily by performing motor task events, it is referred to as event-related desynchronization (ERD) and event-related synchronization (ERS). Importantly, not only motor action but also the sole imagination of movement may evoke an ERD/ERS pattern. Thus, this signal is viable for BCI control.

   Due to the low spatial resolution of the EEG only a limited number of imaginations (referred to as classes) can be distinguished. Depending on a user's individual ERD/ERS

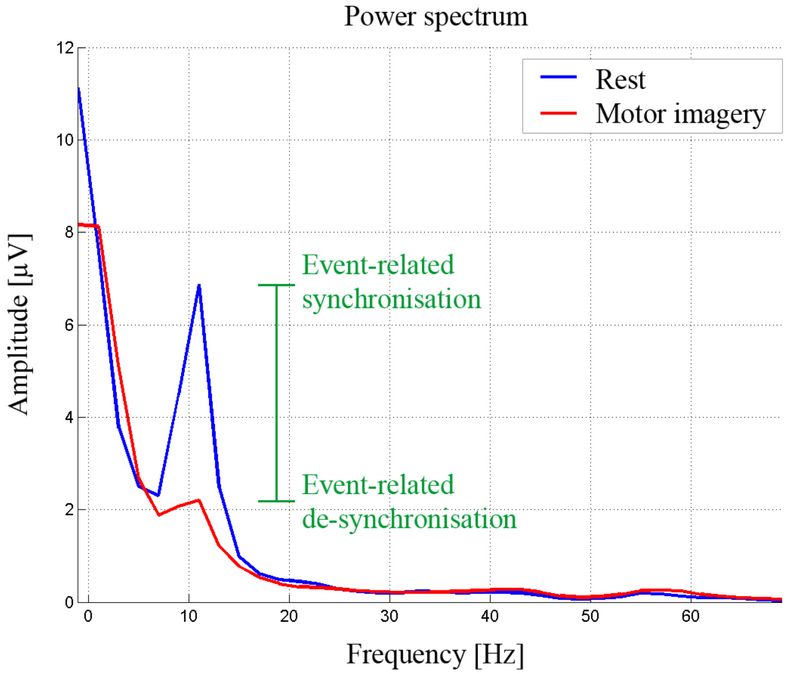


Figure 5: Power spectrum of an EEG recorded over sensorimotor areas from a healthy participant. During rest, the sensorimotor rhythms are highly snychronized (event-related synchronisation, i.e. high power of the SMR). However, once a movement is imagined, the signal de-sychronizes (event-related desynchronisation, i.e. low power of the SMR). The differences in the spectrum can be classified to identify which of both tasks a user is currently performing. Data were obtained in a pilot study during the period of this dissertation.

pattern it is decided how many classes can be trained for BCI control. Usually, two classes work most reliably, for example imagination of grasping movement of one hand against the other or one hand against both feet. In some participants, even three classes may lead to successful control (left hand vs. right hand vs. both feet). Although some users modulate SMR at high accuracy levels it was found that most participants do not achieve accurate control or display large performance variations across sessions and runs (Guger et al., 2009; Blankertz et al., 2010; Halder et al., 2011; Hammer et al., 2012). In a recent study by Holz and colleagues (subm.), a patient achieved high accuracy in a copy-task, however, was not able to use the BCI for sufficient control in a free-mode, potentially due to an increased workload in the online setting. Guger and colleagues (2009) estimated the number of those not able to control a BCI by means of SMR modulation to be 30% and those able to achieve high accuracy to be only 19%. Comparison to an

ERP-BCI system (used in this thesis, see section 1.3.2) revealed that only 3% of users were not able to control such ERP-BCI and 89% of users were able to perform at high accuracy. The term BCI illiteracy has been introduced to display a user's inability to control a BCI, however, as overcoming such low accuracy should primarily be a task for the BCI developer (filters, classifiers, user-friendly feedback, etc.) and only partly for the user (training) it has been proposed to talk of BCI inefficiency instead (Kübler et al., 2011).

SMR-BCIs were proposed for different fields such as (1) communication in healthy (e.g. Millán and Mouriño, 2003; Müller and Blankertz, 2006) as well as those with neurodegenerative disease (e.g., Neuper et al., 2003; Kübler et al., 2005a) (2) control of prosthetic devices (e.g., Pfurtscheller et al., 2000, 2003; Müller-Putz et al., 2005), robots (e.g. Millán et al., 2004) or wheelchairs (e.g., Galán et al., 2008) (3) and recently as a training tool for stroke rehabilitation (e.g., Pichiorri et al., 2011; Grosse-Wentrup et al., 2011).

3. *Steady-state evoked potentials (SSEP)*: SSEPs are oscillatory potential fluctuations triggered by frequently presented, external stimuli and can be recorded in the visual domain (steady-state visual evoked potential, SSVEP), somatosensory domain (steady-state somatosensory evoked potential, SSSEP) and auditory domain (steady-state auditory evoked potentials, SSAEP); for a recent review, e.g. Allison et al., 2012b. A BCI can be implemented by presenting stimuli with different presentation frequency, e.g. checkerboard stimuli with different flicker frequency in the visual domain. Participants continuously focus on one of the stimuli to select a functionality it represents. For example in a navigation task, four stimuli can represent navigation directions and by focusing on one stimulus, participants select a navigation direction (Martinez et al., 2007).

1.3.2 BCIs based on event-related potentials (ERP)

Event-related potentials (ERPs) are positive or negative potential deflections in the EEG that occur time-locked after a known stimulus (figure 6; for review see e.g., Fabiani et al., 2007; Luck, 2005). They are elicited in an oddball paradigm, a procedure in which frequently presented irrelevant stimuli are rarely interspersed with relevant target stimuli (referred to as odd stimuli due to its rarity of occurrence). Focusing attention on the rare target stimuli and ignoring all other non-target stimuli will elicit a cascade of ERPs in response to target stimuli that can be detected from the EEG post stimulus. These potentials are highly reliable in terms of latencies. Thus, once an ERP cascade is detected, the system can easily determine which stimulus led to its elicitation. For example, if one knows that a distinct ERP occurs with a subject-specific latency of 340ms after a stimulus was presented, one can determine which stimulus was presented around 340ms earlier once the distinct ERP was detected from the subject's ongoing EEG. Consequently, it is possible to detect from a set of stimuli which stimulus a user focused attention on (Farwell and Donchin, 1988). This type of BCI has often been referred to as P300-BCI, as it mainly relies on a positive potential called P300 (Sutton et al., 1965). It occurs approximately 300ms post-stimulus but its latency may vary between 200 and 500ms depending on paradigmatic settings (for review, e.g. Polich, 2007). As it is not only the P300 but also other ERPs that contribute to classification accuracy the term ERP-BCI is considered more precise and thus used in this work (e.g., Kaufmann et al., 2011a for analysis of ERPs contributing to classification accuracy across N=51 participants). For a recent review on ERP-BCIs, see e.g., Kleih et al., 2011; Mak et al., 2011; Sellers et al., 2012.

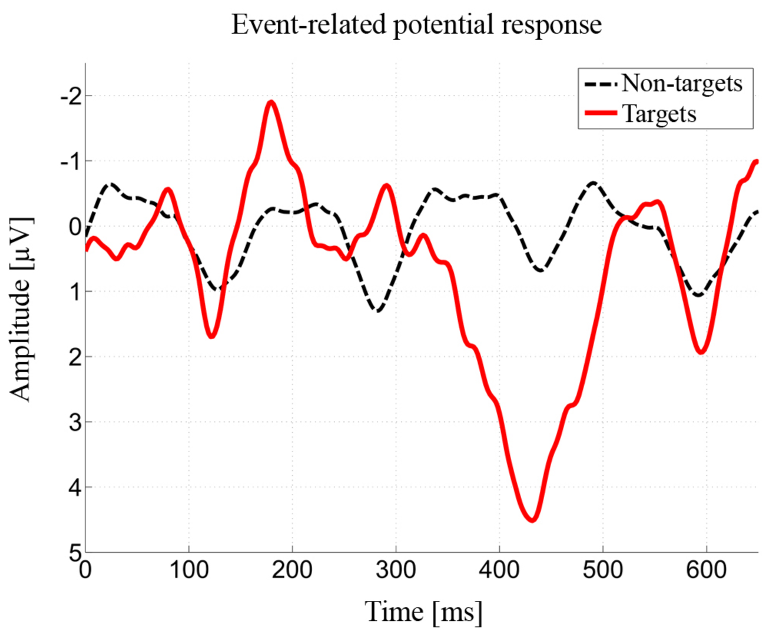


Figure 6: Event-related potential response when attending to target stimuli (solid, red curve) while ignoring all other stimuli (non-targets; dashed, black curve). Around 200ms a negative deflection (N200) is visible followed by a pronounced P300 response peaking around 420 ms post-stimulus.

In a typical ERP-BCI paradigm used for communication, numbers and letters from the alphabet are arranged in a visually displayed matrix (Farwell and Donchin, 1988). An oddball paradigm is realized by highlighting characters randomly. Users focus their attention on one character and count the number of target character intensifications while ignoring flashing of all other characters. As described above, target character-flashes will elicit distinct ERP responses that can be detected from the ongoing EEG recordings. Consequently, it can be determined which character intensification led to elicitation of the detected ERPs, i.e. the intended target character can be determined. Therefore, it is possible to spell on a character-by-character basis. As flashing of single characters in the matrix is time consuming considering the high number of letters in the alphabet, groups of characters are usually flashed at the same time. Typically, rows and columns are flashed randomly (figure 7a, 7b). By detecting which row and which column flash led to an ERP elicitation, the target character can be determined (i.e. the character at the intersection of the detected row and column, see figure 7c for illustration).

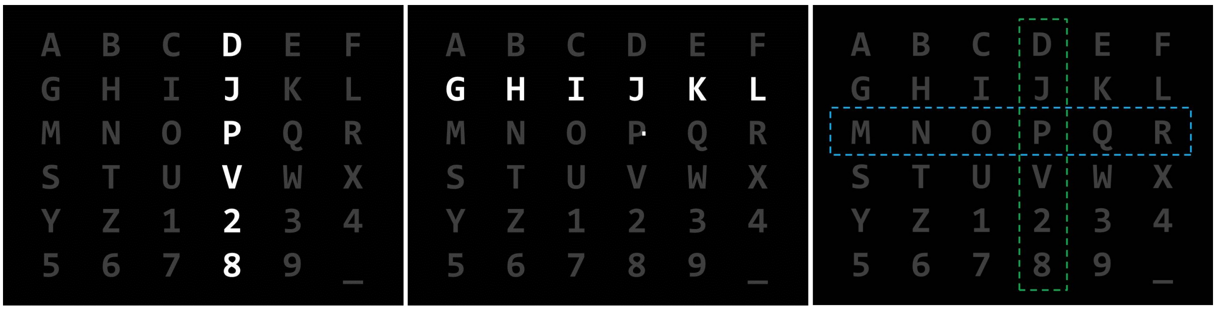


Figure 7: In a typical ERP-BCI paradigm characters are arranged in a matrix and rows and columns are highlighted randomly. If for example P is the intended letter, users focus their attention on the P and count the number of target character-highlightings. (A) Displays a target character-flash and (B) displays flash of non-target characters. (C) By detecting target row (blue line in this case) and target column (green line in this case), it is possible to identify the intended target character at the intersection of the specific rows and columns. In the case of this example, the selected character would be P.

An ERP-BCI session usually comprises three steps. (1) First, a calibration run is performed, in which users are required to spell some characters (usually 8-10 characters) without any feedback provided. This procedure is thus referred to as ”offline”, as no online classification of ERPs is involved. During a calibration run, data is only acquired but not processed. (2) This data is then used to set up a classifier specifically adjusted to a user’s individual ERP response. Different classifiers have been proposed and validated in the literature, among which stepwise linear discriminant analysis (SWLDA) was found superior and is thus commonly used for classification (Farwell and Donchin, 1988; Krusienski et al., 2006; for review, e.g. Kleih et al., 2011). For each electrode, EEG signals post stimulus are analyzed and determination between target and non-target signals is investigated for each time window (usually windows of 2-3ms, referred to as bins, for 800ms post-stimulus). Those bins that contribute best are taken into account for classification. Therefore, a classifier matrix is built that contains a list of electrode numbers, time bins and its corresponding negative or positive classifier weights. (3) During online runs, this matrix then determines which signals are taken into account for classification. After a preset number of stimulation cycles, classification is performed online and immediate feedback is provided to the user, i.e. each selected character is directly displayed so that it is possible to correct for selection errors. An ERP-BCI session usually comprises three steps. (1) First, a calibration run is performed, in which users are required to spell some characters (usually 8-10 characters) without any feedback provided. This procedure is thus referred to as ”offline”, as no online classification of ERPs is involved. During a calibration run, data is only acquired but not processed. (2) This data is then used to set up a classifier specifically adjusted to a user’s individual ERP response. Different classifiers have been proposed and validated in the literature, among which stepwise linear discriminant analysis (SWLDA) was found superior and is thus commonly used for classification (Farwell and Donchin, 1988; Krusienski et al., 2006; for review, e.g. Kleih et al., 2011). For each electrode, EEG signals post stimulus are analyzed and determination between target and non-target signals is investigated for each time window (usually windows of 2-3ms, referred to as bins, for 800ms post-stimulus). Those bins that contribute best are taken into account for classification. Therefore, a classifier matrix is built that contains a list of electrode numbers, time bins and its corresponding negative or positive classifier weights. (3) During online runs, this matrix then determines which signals are taken into account for classification. After a preset number of stimulation cycles, classification is performed online and immediate feedback is provided to the user, i.e. each selected character is directly displayed so that it is possible to correct for selection errors.

Since the introduction by Farwell and Donchin (1988), ERP-BCIs have been intensively used in communication settings in healthy as well as severely impaired patients (e.g., Sellers and Donchin, 2006; Kübler and Birbaumer, 2008; Hoffmann et al., 2008; Nijboer et al., 2008b; Sellers et al., 2010; for review on all ERP-BCI studies including patients until 2010, Mak et al., 2011). Apart from communication, other applications have been suggested. Letters and numbers in the above-described matrix can easily be exchanged to represent functions of assistive technology devices, e.g. for environmental control, internet and email (Zickler et al., 2011). Münßinger and colleagues (2010) suggested an entertainment application for painting. Different painting functions such as color selection, brush size and brush position are placed in a matrix and can be selected in the same manner as described above for communication applications. It has been shown that healthy users as well as severely impaired patients were able to draw predefined paintings or to freely express their creativity (e.g., Holz et al., 2013). Application control can also be achieved without matrix-based stimulation. Zickler and colleagues (2011) validated a system in which small dots were used as stimuli. They were directly presented on top of application functions. For example in an email application, stimuli were directly presented on top of the buttons and users focused attention directly on stimulation of the ”send” button to send an email. ERP-BCIs have also been suggested for wheelchair control, i.e. users select different navigation control options from a visual display mounted on the wheelchair (e.g., Rebsamen et al., 2010; Pires et al., 2008; Iturrate et al., 2009). The proposed systems vary with regard to the amount of control that is left to the user. Some systems propose to select a targeted destination and dedicate navigation control to a smart wheelchair while other systems leave navigation control on the user side.

Apart from the above-described ERP-BCIs that work in the visual modality, non-visual ERP-BCIs were proposed for those not able to perceive visual stimulation or for settings in which visual stimulation is not viable. For example, when considering ERP-BCI based wheelchair control, visual stimulation is not an option as focus on a screen could hamper observation of the environment. In addition, bright sun might complicate visual perception.

Non-visual ERP-BCIs draw on either the auditory or the tactile modality. In line with visual ERP-BCIs they utilize an oddball paradigm for elicitation of ERPs. However, instead of visually presenting stimuli, either auditory (tone) stimuli (e.g., Hill et al., 2005; Sellers and Donchin, 2006; Furdea et al., 2009; Kübler et al., 2009; Klobassa et al., 2009; Halder et al., 2010; Höhne et al., 2010; Schreuder et al., 2010, 2011b; Höhne et al., 2011; Käthner et al., 2013; Hill and Schölkopf, 2012) or tactile (vibration pulses) stimuli (e.g., Aloise et al., 2007; Brouwer and van Erp, 2010; Brouwer et al., 2010; Thurlings et al., 2012; van der Waal et al., 2012) are presented. In a comparison study, Aloise and colleagues (2007) reported that the visual modality usually results in best classification accuracy. However, choice of modality is restricted to the patient’s individual abilities and may well differ from case to case (Kaufmann et al., 2013a).

1.4 A model of BCI control

As described in section 1.3 various types of brain computer interfaces differ with regard to the amount of users able to control them with high accuracy. Inter-subject differences as well as intra-individual differences between trials and across sessions are common.

To explain the variance in BCI performance, Kübler and colleagues (2011) introduced a model of BCI-Control. It summarizes different aspects affecting BCI control that can be grouped into four categories.

1. *Individual characteristics*: On the user side, these are anatomical, physiological and psychological differences. Usually, performance is decreased in patients with neurodegenerative disease, potentially due to degeneration of neurons in motor areas and in the frontal lobes (see section 1.1, e.g., Kübler et al., 2004; Piccione et al., 2006) and a higher training effort (e.g., Kübler et al., 2001b). To estimate the effects of psychological and physiological inter-individual differences, different variables were in the focus of investigation and related to success in BCI tasks. For example, Blankertz and colleagues (2010) identified the mu-peak at rest being a predictor of success in SMR-BCIs, i.e. when recording baseline EEG for 2 min and computing the maximum difference between the power-spectrum density of the EEG and the 1/f decrease, 28% of the variance in BCI performance could be explained. The quick estimation procedure makes such a predictor practically feasible. Halder and colleagues (2011) compared well performing participants to those not able to control a BCI at high accuracy and found differences in the activation patterns of prefrontal areas rendered visible by means of fMRI. Psychological variables were found to moderately affect BCI performance (e.g., Kleih et al., 2010; Nijboer et al., 2010; Hammer et al., 2012; Käthner et al., 2013).

2. *Characteristics of the BCI*: On the technical side, these are differences in software and hardware setup of the system. A plethora of feature extraction and feature translation methods have been introduced and validated (for review, see e.g., Krusienski et al., 2012; McFarland and Krusienski, 2012). Hardware settings as for example different types of electrodes (wet/dry; passive/active; for review e.g., Mak et al., 2011), different number of electrodes at different locations (e.g., Krusienski et al., 2008) or different amplifier settings have been used (for review, see e.g., Wilson et al., 2012).

3. *Feedback and instruction*: Instruction of users and the nature of the feedback provided may affect the performance they can achieve. Neuper and colleagues (2005) found that kinesthetic imagination of movement during an SMR-BCI task yielded better performance than visual imagination, which is due to thereby induced corticomotor excitability (Stinear et al., 2006). Pichiorri and colleagues (2011) further contributed with the finding that goal-oriented imagery significantly increased motor-cortical excitability. Feedback is required to:

   (a) passively present the outcome of classification, i.e. to inform users and allow for their reaction. Nijboer and colleagues (2008a) compared use of visual and auditory feedback; others explored the tactile modality (e.g., Cincotti et al., 2007 for SMR-BCI; Schreuder et al., 2012 for ERP-BCI).

   (b) to actively induce brain responses in the user (e.g. a stimulus in ERP-BCIs). For example multiple modifications regarding the number of targets (Sellers et al., 2006), size, color and distance of characters (Salvaris and Sepulveda, 2009) and stimulus timing (Sellers et al., 2006) as well as different stimulus patterns (e.g. Townsend et al., 2010) have been suggested for ERP-BCIs (for review, see e.g., Sellers et al., 2012).

4. *Application*: Many different applications have been adopted to BCI control. In general, the more complex an application, the more difficult it is to control with a BCI and stepwise increasing difficulty of the task may be required to account for individual difficulties of impaired users as compared to healthy participants in a laboratory environment (Kübler et al., 2001b).

Figure 8 visualizes the four components of the model. The thesis at hand addressed several aspects of this model, suggesting modifications or investigating relations, thereby aiming at improving the outcome of BCI control. The next chapter describes the concept of this thesis and the conducted studies.

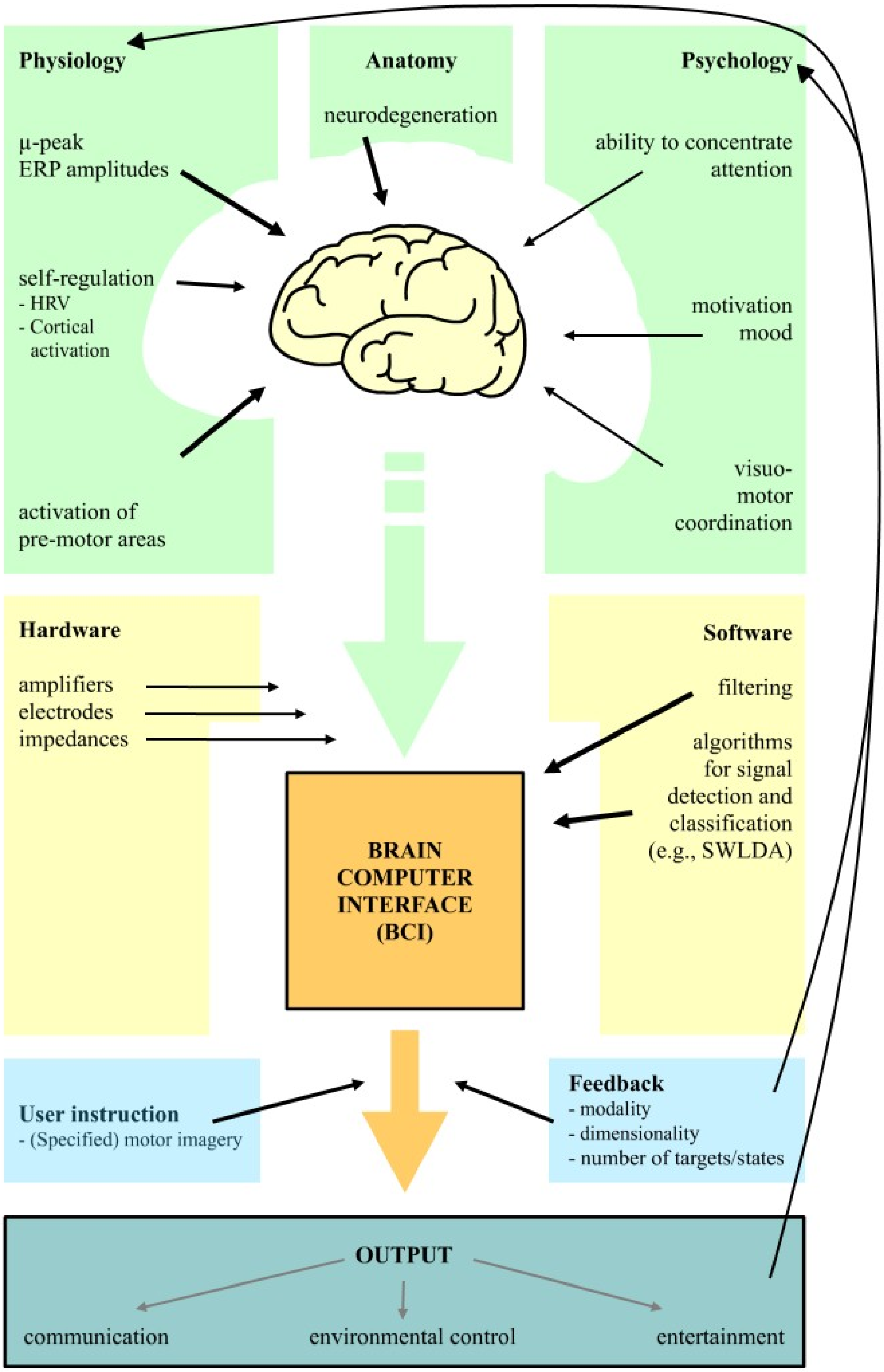


Figure 8: A model of BCI control as published by Kübler and colleagues (2011). The model comprises four categories, i.e. (1) individual characteristics of the user, (2) characteristics of the BCI at hand, (3) feedback and instruction and (4) application output.

1.5 Studies of this dissertation

The main purpose of this thesis was to improve communication based on brain-computer interfaces and thereby enhancing its practical value for people with neurodegenerative disease or brain damage.

As described in detail in section 1.3.1, EEG was chosen as the recording technique of brain signals. Due to its high temporal resolution and low risk, EEG is qualified for BCI-based communication devices, i.e. fast commands can be delivered by means of EEG. Its convenient portability, straightforward setup and comparably low costs render EEG of practical value to be established for daily life communication at potential end-user's homes. Although various EEG based BCI systems have been proposed for communication, this work focused on one EEG component only, i.e. event-related potentials. As previously described in detail (section 1.3.1), SCP-based communication devices are limited in speed and require intensive training (Kübler et al., 1999). SMR-based BCIs are limited in the number of classes that can be distinguished and lack of sufficient reliability in most users (e.g., Guger et al., 2009). SSVEP-based BCIs reliably achieve high accuracies, yet the flicker stimuli may annoy users, in particular reported by elderly persons (Allison et al., 2010). The number of classes in SSEP-based BCIs are not as limited as SMR-based BCIs, yet a large number of characters to be selected by a user for communication purpose can most easily be handled with ERP-based BCIs (see section 1.3.2). ERP-BCIs were found to be highly reliable (Guger et al., 2009) and stable over time (Sellers et al., 2010; Nijboer et al., 2010; Holz et al., 2013) and have already been successfully used as a communication tool for severely impaired patients (for review, Mak et al., 2011).

However, the practical value of BCI systems as a means of communication is still low. Developing systems that are clinically useful is an inevitable requirement for the future of the field (e.g., Zickler et al., 2011; Wolpaw and Wolpaw, 2012b; Vaughan et al., 2012). In the planning phase of this thesis project, basic issues were thus identified that remained to be addressed:

1. **Can we identify performance predictors so that we can provide users with individual BCI solutions without the need of multiple, demanding testing sessions?**

   If one could reliably estimate a user's performance with different BCI systems without multiple testing sessions but with a short assessment of one or several performance predictors, this would increase the practical value of BCI systems. Several predictors have already been identified for different BCI systems, yet a more global picture is still needed (e.g., Kleih et al., 2010; Blankertz et al., 2010; Halder et al., 2011; Hammer et al., 2012; Halder et al., 2013). Furthermore, such predictors contribute to our knowledge on factors influencing BCI control, thereby helping to identify sources of inter- and intra-individual performance differences. Finally, identified predictors could result in specific training of the user, i.e. improving a factor that is found to strongly influence BCI control abilities (Botrel et al., 2013).

2. **Can we provide complex BCI technology in an automated, user-friendly and easy-to-use manner, so that BCIs can be used without expert support at end-users' homes?**

   Zickler and colleagues (2009) validated the requirements of new technology in N=77 assistive technology (AT) users. Although generally satisfied with their current AT solutions users reported a need for better technology. Functionality, easiness of use and the possibility of independent use were rated as most important. In a follow-up study, N=4 potential end-users validated current BCI technology (Zickler et al., 2011). None of them would use the technology in its current state of development if they had alternatives. Users negatively rated (1) high EEG preparation effort, (2) complexity of hard- and software setup and (3) the low communication speed compared to existing AT solutions.

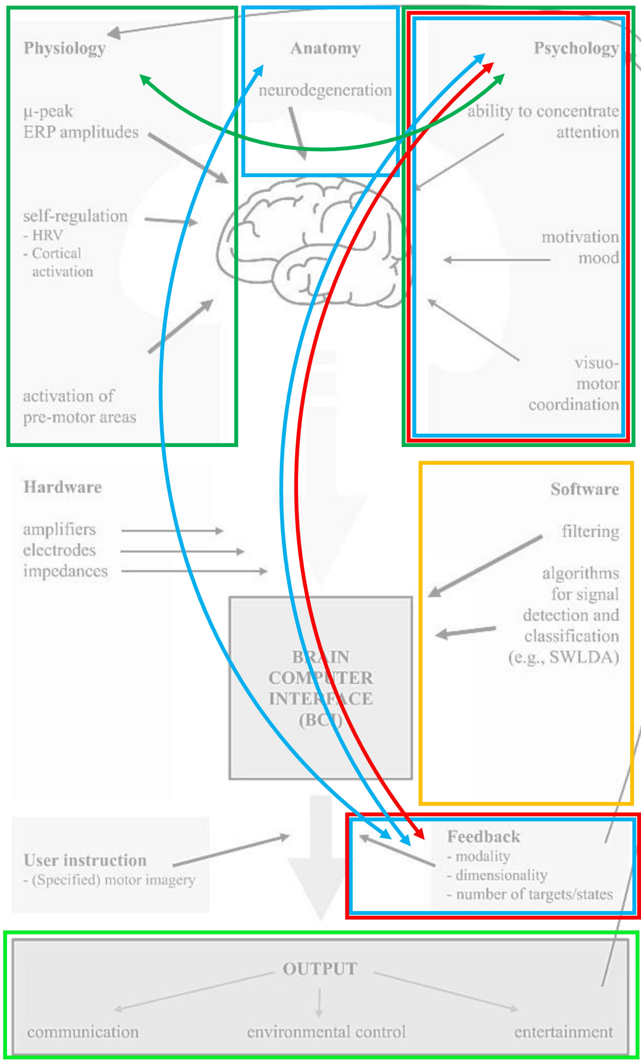
3. **How can we account for and improve the low information transfer rates as compared to other existing assistive technology solutions?**

   With the current state of technology, BCIs are only of value for a very small population of potential users with the severest disabilities (for review, e.g., Huggins and Zeitlin, 2012), i.e. people with severe LIS (see chapter 1.1). For most users depending on assistive communication tools, other options provide higher information transfer rates (i.e. the amount of information correctly transmitted per time unit) and may well be preferred, e.g. when capable of using an eye-tracking device (Cipresso et al., 2012). BCI systems thus have to improve efficiency (the information transfer rate) to achieve wider acceptance (Huggins and Zeitlin, 2012; Zickler et al., 2011).

4. **How can we prevent the performance drop often seen when bringing BCI technology that was tested in healthy users to those with severe motor impairment?**

   Initial testings of new developments and basic research investigations are usually performed in healthy samples to reduce patient effort. Healthy participants can provide a first estimation of feasibility. Yet, testing BCI technology in healthy users only, and often in rather small samples is a common phenomenon in the field of BCI (e.g., Holz et al., 2012; Vaughan et al., 2012). This is a critical lack of validation in particular as performance usually decreases when transferring the technology to patients (Piccione et al., 2006). Research should thus investigate how to prevent such a performance drop, i.e. identifying its sources and suggesting changes to overcome this BCI inefficiency phenomenon (see chapter 1.3.1).

To address these issues, the work at hand targeted several aspects of the model suggested by Kübler and colleagues (2011, see chapter 1.4). Figure 9 illustrates these aspects and their interconnections.



Can we identify performance predictors so that we can provide users with individual BCI solutions without the need of manifold, demanding testing sessions?

How can we prevent the performance drop often seen when bringing BCI technology that was tested in healthy users to those with severe motor impairment?

Can we provide complex BCI technology in an automated, user-friendly and easy-to-use manner, so that BCIs can be used without expert support at end-users homes?

Measurements of outcome: Predictive value of assessed variables, achieved spelling accuracy, achieved bit rate, impact on BCI inefficiency, questionnaires, ERP amplitudes and latencies, determination coefficients, spelling times, offline classification accuracies, etc.

How can we account for and improve the low information transfer rates as compared to already existing assistive technology solutions?

Figure 9: Basic issues identified for this thesis project and their interconnection based on the model of BCI control by Kübler and colleagues, 2011.

In the following, a synopsis of the studies conducted to address the four questions raised is provided:

1. **Can we identify performance predictors so that we can provide users with individual BCI solutions without the need of multiple, demanding testing sessions?**

   A spelling task in an ERP-based BCI requires users to sustain their attention on stimulations of the intended target character while ignoring stimulations of any other character (see chapter 1.3.2). This requires inhibitory control capability triggered by the prefrontal cortex. In their model of neurovisceral integration, Thayer and Lane (2000) suggested a cortico-cardiac interaction circuit of central to peripheral-physiological regulation. It is quantified by assessing different measures of heart rate variability (HRV). Thus, based on the proposed cortico-cardiac interaction, HRV has been introduced as a trait index of a person's attentional and cognitive resource allocation capabilities (for review, e.g., Thayer and Brosschot, 2005). Decreased HRV is associated with poor attentional control capability (for review, e.g., Thayer and Siegle, 2002).

   The variability of time intervals between heartbeats is modulated by vagal and sympathetic efferent neurons (for review, e.g., Saul, 1990). HRV is quantified by several estimates in the time and frequency domain (TaskForce, 1996; Berntson et al., 1997) that were each differently associated to represent either more vagal or more sympathetic influences. Since strength in vagally-mediated HRV is associated with higher attention allocation and inhibition of irrelevant information (Porges, 1992, p. 208; Thayer and Lane, 2000), this work assessed whether tonic vagal activation could be used as a predictor of ERP-BCI spelling performance. Such predictor would be easy to obtain from a 5-min resting period recording of heartbeats independent from actual BCI use.

   $\hookrightarrow$ A study was conducted investigating predictive value of HRV in N=34 healthy BCI users. Patients were not yet targeted in this basic investigation. The findings were published in the International Journal of Psychophysiology (Kaufmann et al., 2012a) and are included in chapter 2.1.

   Berntson and Stowell (1998) demonstrated that estimates of HRV are particularly prone to artifacts as even single artifacts have severe impact on the computed frequency estimates (figure 10). Missing software solutions for appropriate handling of artifacts in electrocardiogram (ECG) and inter-beat interval (IBI) data, consequently encouraged development of our own tool named *ARTiiFACT*. Its core feature lies in the implementation of an artifact detection method theoretically proposed by Berntson and colleagues (1990). This algorithm derives a data-specific threshold criterion from percentile-based distributions of IBIs. This criterion constitutes the mean of the maximum expected beat difference of valid beats (MED) and the minimum expected difference of artificial beats (MAD):

$$\text{Criterion Beat Difference} = \frac{MAD + MED}{2} \tag{1}$$

Both are derived from the inter-quartile range (IQR) of the data as:

$$MED = 3.32 \times \frac{IQR}{2} \tag{2}$$

$$MAD = \frac{\text{Median Beat} - 2.9 \times \frac{IQR}{2}}{3} \tag{3}$$

Thus, as the criterion is obtained from percentile-based distributions, it is less prone to extreme values as for example a simple mean, i.e. it is less affected by artifacts in the data (Berntson et al., 1990). Finally, the criterion is applied to the data for detecting artificial beats that can then be deleted or interpolated.

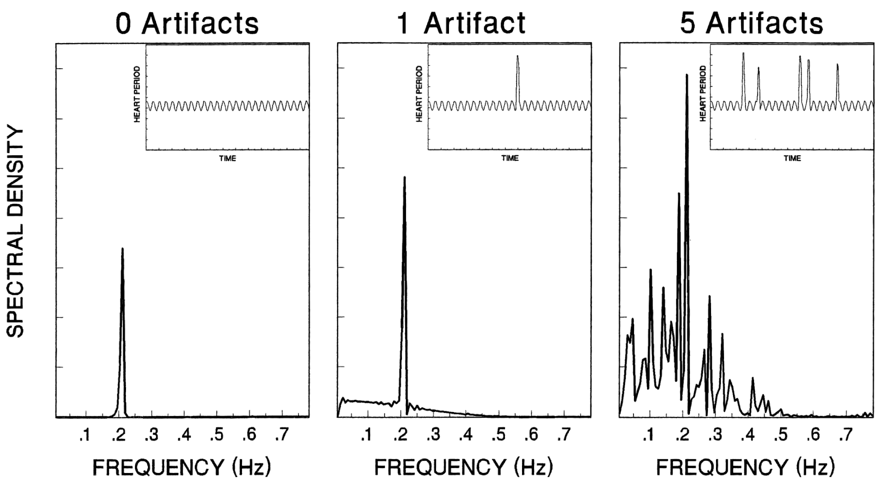


Figure 10: Effect of artifacts on computed heart rate variability. Missed beats result in an increase and blurring of the computed spectral power estimates. Figure reproduced with permission from Berntson and Stowell (1998), Copyright John Wiley and Sons. The publication is available online at http://onlinelibrary.wiley.com.

Implementation of the algorithm was validated against a software named Kubios that can be considered a standard software for HRV analysis (Niskanen et al., 2004; Tarvainen et al., 2009). It provides artifact detection at five levels of sensitivity. For systematic validation, artificial data sets were automatically generated. Therein induced artifacts systematically varied with regard to how pronounced and of which type (missed-beat or falsely detected beat) they were. As such, it was possible to systematically reveal strengths and weaknesses of the detection methods.

ARTiiFACT provides a graphical user interface and is free available upon request. It allows conducting the full cascade of ECG data analysis including:

- Preprocessing of ECG raw data, detection of heartbeats and extraction of inter-beat intervals (IBI)
- Detection and processing of artifacts in IBI data (see above)
- Computation of HRV parameters in both time and frequency domain
- Statistical validation of data quality

↪ Results of the validation study as well as a detailed description of the software were published in Behavior Research Methods (Kaufmann et al., 2011c). The publication is included in chapter 2.2 of this thesis.

2. **Can we provide complex BCI technology in an automated, user-friendly and easy-to-use manner, so that BCIs can be used without expert support at end-users' homes?**

   As described earlier in this chapter, the work of Zickler and colleagues (2009; 2011) pointed out that potential BCI end-users request a technology that is easy-to-use and can be used independently from professionals. Thus, a study was conducted to assess if current BCI software technology can be handled in an automated manner, so that the required level of device control is limited to a start/stop operation only. To achieve this, a commonly used BCI software (BCI2000, Schalk et al., 2004) was modified so that it sets all necessary parameters automatically. Calibration data is analyzed using MATLAB (The Mathworks Inc., USA) - also automatically and without the requirement of a BCI expert to control the data, - classifier weights are computed and fed back into the BCI system. None of these processes are visible to the person controlling the BCI.

↪ In a validation study, N=19 BCI novices controlled the software on their own and without expert support (apart from EEG setup and instruction prior to BCI use). Results of this study and details on the software implementation were published in Frontiers in Neuroprosthetics (Kaufmann et al., 2012b). The publication is included in chapter 2.3.

3. **How can we account for and improve the low information transfer rates as compared to other existing assistive technology solutions?**

   In Kaufmann and colleagues (2012b, chapter 2.3), a predictive text entry (PTE) system was directly integrated into the spelling matrix. Predictive text suggestions were presented in a separate column and could be selected in the exact same manner as characters, i.e. by focusing attention on stimulation of the targeted text suggestion. As performance did not decrease compared to spelling without PTE, this is a valuable way to increase information transfer rates.

   Yet another approach to address this issue is to improve the quality of the underlying EEG signal, thereby enhancing performance of the BCI classification algorithm. The commonly used ERP-BCI approach, in which rows and columns of a matrix are highlighted randomly (Farwell and Donchin, 1988), usually requires multiple repetitions of the matrix stimulation cycle until sufficient classification certainty can be achieved. Improving the signal quality may promote improved classification accuracy with fewer repetitions.

   To achieve this, the BCI paradigm was modified such that it overlays characters with pictures of famous faces instead of simply highlighting them. Faces elicit a prominent cascade of event-related potentials involved in their perception and processing. Apart from the P300, these are the N170 (around 170 ms; Bentin et al., 1996) and N400f (between 300 and 500 ms; Eimer, 2000). The N170 is involved in rapid perception of faces whereas the N400f is involved in processing of familiar faces (Eimer, 2000). Consequently, with face stimuli the signal-to-noise ratio of the EEG post-stimulus could be increased compared to simple character-highlighting.

   ↪ This new paradigm was validated in an offline setting with N=20 healthy participants comparing it against the commonly used paradigm. Results of this study were published in the Journal of Neural Engineering, including an analysis of the elicited ERP responses and a comparison of achieved offline classification estimates (Kaufmann et al., 2011b).

The publication is included in chapter 2.4.

4. **How can we prevent the performance drop often seen when bringing BCI technology that was tested in healthy users to those with severe motor impairment?**

   As the offline validation results of the new face paradigm were promising (Kaufmann et al., 2011c), the paradigm was brought to N=9 severely impaired patients with neurodegenerative disease. In an online setting, patients spelled several words while the number of stimulation cycle repetitions was step-by-step decreased. This allowed for a close investigation of a potential benefit of the new paradigm on online information transfer rates. Furthermore, a comparison of EEG signals acquired during use of the classic BCI paradigm and during use of the face paradigm was conducted. Performance was compared to N=16 healthy participants to investigate the impact of the new paradigm on the often reported performance drop in patients (Piccione et al., 2006). Furthermore, it was investigated how variations of the stimulus material in terms of familiarity of the presented face affects the ERPs and thereby affects performance.

   ↪ Results of this study with N=9 patients and N=16 healthy participants were published in Clinical Neurophysiology (Kaufmann et al., 2013b) and are included in chapter 2.5 of this work.

2 Publications

2.1 Kaufmann, T., Vögele, C., Sütterlin, S., Lukito, S., and Kübler, A. (2012). Effects of resting heart rate variability on performance in the P300 brain-computer interface. International Journal of Psychophysiology

Abstract: *Objective*: Brain computer interfaces (BCI) can serve as a communication system for people with severe impairment in speech and motor function due to neurodegenerative disease or injury. Reasons for inter-individual differences in capability of BCI usage are not yet fully understood. Paradigms making use of the P300 event-related potential are widely used. Success in a P300 based BCI requires the capability to focus attention and inhibit interference by distracting irrelevant stimuli. Such inhibitory control has been closely linked to peripheral physiological parameters, such as heart rate variability (HRV). The present study investigated the association between resting HRV and performance in the P300-BCI. *Methods*: Heart rate was recorded from 34 healthy participants under resting conditions, and subsequently a P300-BCI task was performed. *Results*: Frequency domain measures of HRV were significantly associated with BCI-performance, in that higher vagal activation was related to better BCI-performance. *Conclusions*: Resting HRV accounted for almost 26% of the variance of BCI performance and may, therefore, serve as a predictor for the capacity to control a P300 oddball based BCI. *Significance*: This is the first study to demonstrate resting vagal-cardiac activation to predict capability of P300-BCI usage.

Link to publication source: International Journal of Psychophysiology, Elsevier B.V.

International Journal of Psychophysiology 83 (2012) 336–341

Contents lists available at SciVerse ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

Effects of resting heart rate variability on performance in the P300 brain-computer interface

Tobias Kaufmann [a,*], Claus Vögele [b], Stefan Sütterlin [b,c], Steve Lukito [d], Andrea Kübler [a,e]

[a] Department of Psychology I, University of Würzburg, Marcustr. 9-11, D-97070 Würzburg, Germany
[b] Research Unit INSIDE, University of Luxembourg, Route de Diekirch, L-7220 Walferdange, Luxembourg
[c] Research Group on Health Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium
[d] Graduate Training Centre of Neuroscience, International Max Planck Research School, University of Tübingen, Österbergstr. 3, D-72074 Tübingen, Germany
[e] Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Gartenstr. 29, D-72074 Tübingen, Germany

ARTICLE INFO

Article history:
Received 29 June 2011
Received in revised form 28 November 2011
Accepted 29 November 2011
Available online 13 December 2011

Keywords:
Brain-computer interface
P300-BCI
Cardiac autonomic regulation
Heart rate variability
Executive function

ABSTRACT

Objective: Brain computer interfaces (BCI) can serve as a communication system for people with severe impairment in speech and motor function due to neurodegenerative disease or injury. Reasons for inter-individual differences in capability of BCI usage are not yet fully understood. Paradigms making use of the P300 event-related potential are widely used. Success in a P300 based BCI requires the capability to focus attention and inhibit interference by distracting irrelevant stimuli. Such inhibitory control has been closely linked to peripheral physiological parameters, such as heart rate variability (HRV). The present study investigated the association between resting HRV and performance in the P300-BCI.
Methods: Heart rate was recorded from 34 healthy participants under resting conditions, and subsequently a P300-BCI task was performed.
Results: Frequency domain measures of HRV were significantly associated with BCI-performance, in that higher vagal activation was related to better BCI-performance.
Conclusions: Resting HRV accounted for almost 26% of the variance of BCI performance and may, therefore, serve as a predictor for the capacity to control a P300 oddball based BCI.
Significance: This is the first study to demonstrate resting vagal-cardiac activation to predict capability of P300-BCI usage.

1. Introduction

Brain-computer interfaces (BCI) utilize neurophysiological signals, originating in the brain, to control external devices and computer applications (for review, see Birbaumer and Cohen, 2007; Kübler et al., 2001a). In 1988, Farwell and Donchin introduced a BCI based on event-related potentials (ERP) which offers a communication channel independent from voluntary muscular control (Farwell and Donchin, 1988; for review, see Kleih et al., 2011). Such BCI is often referred to as P300-BCI, as it mainly utilizes the P300 event related potential (e.g. Polich et al., 1997; Sutton et al., 1965, for review see Picton, 1992 and Polich, 2007), a prominent, positive deflection in the event related electroencephalogram (EEG). The P300 is evoked by the random and rare presentation of target stimuli that require updating of current memory traces (Donchin, 1981; Verleger, 1988). Depending on the complexity of the stimulus its latency varies from 200 ms to 700 ms. The P300 can be recorded best over centro-parietal areas (Picton, 1992) but its anatomical sources are often subcortical areas, for example the hippocampus (Halgren et al., 1980). The positive polarity of the P300 indicates an inhibitory function, probably blocking competing information processing in the presence of new and challenging material.

The P300-BCI involves an oddball paradigm, in which a rarely presented target stimulus (the *odd* stimulus) is attended to among a plethora of unattended stimuli. For the purpose of communication, users are presented with a matrix consisting of numbers and letters from the alphabet (see Fig. 1). These characters are light flashed row- and column-wise in random order (Farwell and Donchin, 1988). To communicate a character, users focus their attention on the intended character (target character) by mentally counting how often it flashes. As the occurrence of target flashing is rare compared to flashing of all other characters, event-related potentials (particularly the P300) are elicited post-target stimulus. By classification of such ERPs, conclusion can be drawn about the intended character. Hence, the P300-BCI allows users to communicate by selecting characters from the matrix.

In contrast to many other existing BCIs based on neurophysiological signals, such as e.g. slow cortical potentials or sensory motor rhythms, the P300-BCI does not depend on learning to self-regulate the brain response. The short latency of the P300 allows for fast character selection, which is desirable when aiming at daily life application for BCIs. Most

* Corresponding author. Tel.: +49 931 31 81084; fax: +49 931 31 82424.
E-mail addresses: tobias.kaufmann@uni-wuerzburg.de, tobiaskaufmann@web.de (T. Kaufmann), Claus.Voegele@uni.lu (C. Vögele), Stefan.Suetterlin@uni.lu (S. Sütterlin), steve.lukito@lincoln.oxon.org (S. Lukito), andrea.kuebler@uni-wuerzburg.de (A. Kübler).

0167-8760/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpsycho.2011.11.018

T. Kaufmann et al. / International Journal of Psychophysiology 83 (2012) 336–341 337

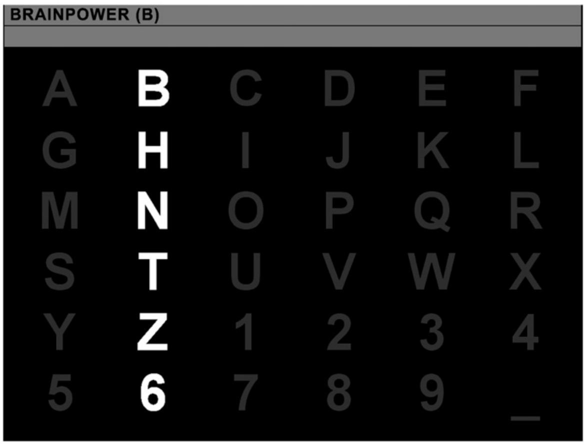


Fig. 1. The visual P300-BCI spelling paradigm. In this snapshot in time, the character B is the target (first letter in the word BRAINPOWER) and the column containing the target is light flashed (odd stimulus).

importantly, the P300-BCI has been shown to be feasible for patients in the locked-in state (e.g. Hoffmann et al., 2008; Kübler et al., 2009; Nijboer et al., 2008; Sellers et al., 2010; Silvoni et al., 2009; for a review, see Mak et al., 2011). The locked-in state refers to a condition of severest motor paralysis with only few muscular movements such as eye blinks or thumb movement, available for communication (Kübler and Birbaumer, 2008).

Nevertheless, interindividual differences between participants' accuracy in BCI usage are common (Blankertz et al., 2010; Guger et al., 2009). For example, a failure to focus eye-gaze was found to negatively influence P300-BCI performance (Brunner et al., 2010; Treder and Blankertz, 2010). Moreover, successful use of the P300-BCI requires the ability to internally spell and intact attentional resources. Such cognitive factors are likely to contribute to individual differences in BCI performance, although systematic investigations are rare (for BCIs based on sensorimotor rhythms, see Hammer et al., 2011). In the P300-BCI spelling paradigm it is often the case that intensified non-targets are located in the central visual field of the user, e.g. non-targets surrounding the target character. Therefore, voluntary focusing of attention and inhibition of task irrelevant information (thus, often referred to as sustained attention), is required.

Inhibitory control has been closely linked to the peripheral physiological parameter heart rate variability (HRV; for definitions and review, see Task Force, 1996; Appelhans and Luecken, 2006). Heart rate variability reflects variations in the timing of consecutive heartbeats that are triggered by input from vagal cardiac and sympathetic efferent neurons (Saul, 1990). These neurons are integrated in the central autonomic network (CAN; Benarroch, 1993, 1997), a functional unit comprising prefrontal and limbic structures, resulting in a direct link between CAN and HRV (Thayer and Lane, 2000). Thayer and colleagues (Thayer and Brosschot, 2005; Thayer and Lane, 2000) have outlined a model of neurovisceral integration, which describes this relationship as a dynamically adaptable circuit system. As such, cardiovascular control is not only restricted to the CAN's output (feed-forward) but also integrates sensory information from the heart that is fed back to the CAN (feedback). Thus, distinct inhibitory processes, originating in the prefrontal cortex modulating sympathetic and vagal cardiac control, play a decisive role in inhibition and regulation of behavior and their failure may cause psychological symptoms such as deficits in attentional or emotional control (Friedman and Thayer, 1998; Thayer and Lane, 2000; Thayer and Siegle, 2002).

Previous studies demonstrated that such inhibitory processes are particularly controlled by the prefrontal cortex (Ahern et al., 2001; for review see Thayer et al., 2009), which fosters involvement of cognitive control. Thus, researchers proposed that HRV, as an index of prefrontal activation, is related to cognitive performance and executive function (Appelhans and Luecken, 2006; Thayer et al., 2009). Experimental data support this notion (for a review, see Thayer et al., 2009). Hansen et al. (2003), e.g. reported on increased cognitive performance particularly for tasks involving inhibition-related executive functions in participants with higher vagal activity at rest, i.e. an index of higher inhibitory capacity. In a similar vein, higher order processes such as overcoming distractory and irrelevant affective information in favor of rational and effective decision-making in cognitively demanding or complex social interaction have also been reported (Sütterlin et al., 2011a, 2011b).

Nevertheless, performance in the tasks employed in these studies on the association of HRV and executive function all involved overt motor responses to target stimuli presented on a screen, e.g. by pressing a designated key on a keyboard. In contrast, performance in the P300-BCI is based on the neurophysiological potentials elicited by an oddball paradigm involving an inhibitory cognitive process, i.e. sustained attention. The P300-BCI, therefore, provides a more direct measure of the inhibitory function that is hypothesized to be indexed by HRV. To the best of our knowledge there is only one study, which investigated the relationship between an oddball paradigm and HRV (Valkonen-Korhonen et al., 2003). The authors, however, investigated changes in HRV in response to mental load (e.g. an oddball task), and not in terms of predicting task performance by resting HRV.

The aim of the present study was to investigate the association between resting HRV and performance in the P300-BCI in healthy participants to provide first results that may be followed up in patient samples compared with an age- matched control group. We hypothesized that HRV would be positively related to P300-BCI performance and may therefore serve as a predictor of BCI performance.

2. Methods

2.1. Participants

Participants were N = 39 university students who reported to be in good health and not to be suffering from any acute or chronic condition. Due to equipment failure or poor compliance with the experimental procedure, data from five individuals had to be discarded. The final sample comprised of N = 34 participants (16 men) with a mean age of 26.3 years (SD = 6.1, range 17–44). The study was approved by the University Ethics Board of Roehampton University of London (UK) and performed in accordance with the Declaration of Helsinki. All participants gave informed consent prior to the experiment.

2.2. Procedure

Participants attended individually for the experimental session. On arrival, the participant was seated in an armchair 1.20 m from the monitor displaying the visual P300 spelling matrix (see Fig. 1). The procedure was explained and electrodes attached. A 10-minute monitored rest period followed (baseline, last 5 minutes were taken into account for analysis), during which participants were asked to sit quietly and relax while beat-by-beat heart rate (HR) was recorded. Then an initial run of the visual P300 spelling matrix was conducted, during which participants were asked to spell the word BRAINPOWER. No feedback was provided in this run. Data were used to train a classifier to enable online feedback-of-results in the following runs. In each of these following runs participants were required to spell one word of five letters each (12 words in total). The runs were interspersed with 30-second breaks (recovery times). After completing the experimental procedure electrodes were removed and participants debriefed.

2.3. Equipment and data acquisition

The electrocardiogram (ECG) was recorded at a sampling rate of 256 Hz from disposable electrodes attached in the lead II configuration and amplified using a Biosemi ActiveTwo AD-box.

Stimulus presentation and data collection were controlled by the BCI2000 software (Schalk et al., 2004; http://www.bci2000.org). Electroencephalogram (EEG) was recorded from 30 Ag-AgCl electrodes at Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, Oz, O2, PO9, and PO10 (according to the extended 10–20 system; guidelines of the American Electroencephalographic Society) and referenced to the average of left/right earlobes. Impedance was kept below 5 kΩ. The EEG was amplified using a 32-channel Biosemi polygraph, sampled at 256 Hz. Fifty-Hertz noise was filtered using the notch filter implemented in the BCI2000 software.

2.4. Visual P300 spelling matrix

The visual P300 spelling matrix was displayed on a 19-inch monitor. It presented a matrix of 6×6 characters (numbers and letters from the English alphabet, see Fig. 1). Each row and column was intensified in random order. During each trial, the participant was asked to focus attention on one of the 36 characters displayed in the matrix, and to count how many times it flashed. The random sequence of six row and six column flashes constitutes an oddball paradigm, with the row and the column containing the target character constituting the rare target stimuli. Each run comprised 8 (sequences)×2 (targets)×5 (letters) = 80 target trials, and 8×10×5 = 400 non-target trials. For a detailed description of the visual P300 speller, see Farwell and Donchin (1988) and Sellers and Donchin (2006).

2.5. Data reduction and analysis

EEG data were filtered between 0.1 Hz (high pass) and 30 Hz (low pass) using Brain Vision Analyzer (Brain Products GmbH, Germany). Data were then segmented in time windows of 800 ms (700 ms post stimulus). Grand average over all target trials and over all non-target trials was calculated across subjects. From Fig. 2 it can be seen that the prominent ERP to target stimuli was the P300 with average peak amplitude of 4.09 μV at 247.7 ms post stimulus. Performance in the BCI was operationalized as percentage of correct copy-spelled letters.

Offline analyses for HRV was performed using *ARTiiFACT* (Kaufmann et al., 2011) and included the extraction of interbeat intervals (IBI) from ECG recordings during baseline. Artifacts were detected via an individually calculated distribution-related threshold criterion (Berntson et al., 1990; Berntson and Stowell, 1998), deleted, and values estimated via linear interpolation of neighboring IBIs (for details please see Kaufmann et al., 2011). Statistical parameters of HRV (Allen et al., 2007; Task Force, 1996) were calculated using the Kubios HRV Analysis 2.0 software (Niskanen et al., 2004). Time domain measures (for an overview: Task Force, 1996) included mean heart rate, RMSSD (square root of the mean squared differences of successive interbeat intervals), and pNN50 (NN50 is the number of interval differences of successive interbeat intervals greater than 50 ms, and pNN50 the proportion derived by dividing NN50 by the total number of interbeat intervals). Spectral frequency measures were derived using Fast Fourier Transformation (FFT). Frequency bands were labeled as recommended by the Task Force (1996) as high-frequency (HF, 0.15–0.4 Hz) and low-frequency (LF, 0.04–0.15 Hz) and expressed in power [ms^2] and percent [%].

Which HRV parameters at best reflect vagal activity as an index of inhibitory functioning is a controversial issue (e.g. Appelhans and Luecken, 2008), as it is barely possible to completely separate vagal from sympathetic activity (e.g. Hedman et al., 1995). Hayano et al. (1990) suggested an attempt to normalize HRV power in the high frequency band (HF-HRV power) with mean interbeat interval. It was found that such RSAnorm (normalized respiratory sinus arrhythmia, also referred to as Hayano index) is much less influenced by sympathetic activity than uncorrected HRV parameters (for a review see Grossman and Taylor, 2007). RSAnorm can be considered an indicator for vagal cardiac activity and thus a physiological marker of inhibitory capacity. Higher values of RSAnorm indicate higher vagal activation.

Q-Q-probability plots were used to test for normal distribution, and not-normally-distributed data (LF/HF) were ln-transformed, which rendered these data normally distributed.

Predicting BCI performance based on physiological parameters requires a robust regression model. We tested the robustness of the association between RSAnorm and BCI performance with a robust regression method (robustfit), pre-implemented in Matlab® (The Mathworks, USA). Computation is performed by iteratively reweighing least squares with a bisquare weighting function until the error is reduced to minimum, thus effectively reducing influence of outliers on the predictive model.

Influence of age and gender was controlled in a partial correlation. Subsequently, influence on correlation coefficient was statistically tested for, using Steiger's Z test for comparing elements of a correlation matrix within a population (Steiger, 1980).

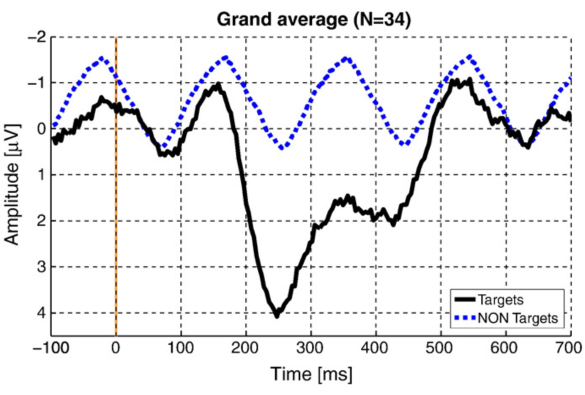


Fig. 2. Grand average of ERPs to targets and non-targets across all participants.

3. Results

Results of HRV analysis were all within normal range (Table 1) and did not show any significant gender effect (all $p > .05$). Also, BCI performance was not significantly different between women ($M = 96.4\%$, $SD = 4.31$) and men ($M = 92.49\%$, $SD = 7.27$; $t(32) = 1.97$, $p > .05$, $d = 0.68$). As we were particularly interested in the association between performance and RSAnorm, we controlled for variance in gender and age but did not find any effect on the reported association (gender: Steiger's $Z = .52$, $df = 31$, $p > .05$; age: Steiger's $Z = .34$, $df = 31$, $p > .05$).

RSAnorm as an index of cardiac vagal activation correlated highly significantly with performance in the P300-BCI ($r = .506$, N = 34,

Table 1
Range, means, and standard deviations of HRV parameters and BCI performance.

	Min	Max	Mean	SD
RMSSD	10.99	88.46	39.58	19.98
pNN50	0	57.19	19.62	17.51
LF Power [%]	9.61	64.75	32.81	12.95
HF Power [%]	4.18	54.23	28.95	13.58
LF/HF	.28	8.82	1.64	1.62
RSAnorm	.72	1.98	1.47	.36
BCI Perform. [%]	73.3	100.0	94.6	6.14

T. Kaufmann et al. / International Journal of Psychophysiology 83 (2012) 336–341 339

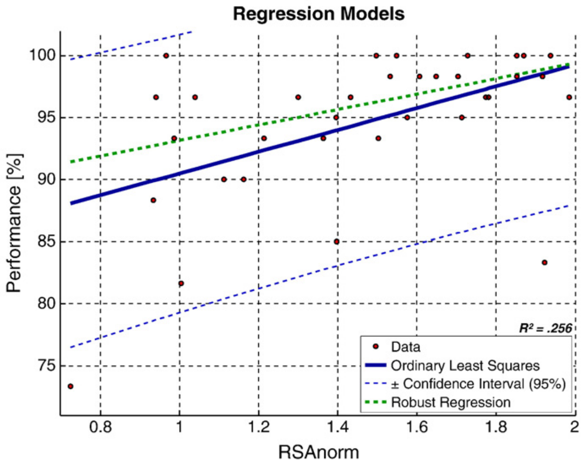


Fig. 3. Scatter-plots displaying the association between BCI performance and RSAnorm, an index of vagal activation. Both, regression of ordinary least squares and robust regression display the same positive association of vagaly mediated HRV with performance in the P300-BCI. Robust regression was used to evaluate the model's sensitivity to outliers. Difference between the models is not significant, therefore confirming stability of the generated model.

$p<.001$). Thus, RSAnorm accounted for 25.6% of the observed variance in BCI performance (see Fig. 3).

This strong relationship between cardiac vagal activation and BCI performance was also reflected in other HRV parameters of the frequency domain. HF-HRV and the ratio of low and high frequency (autonomic balance) correlated highly with BCI performance (all $p<.01$, see Table 2). Importantly, LF-HRV, which is to some extent influenced by sympathetic activity, did not significantly correlate with performance. This further strengthens the hypothesized relationship between vagal activation and BCI performance. Correlation of BCI performance and parameters of the time domain parameters did not reach significance (all $p>.05$, see Table 2).

3.1. Robustness check

The robust regression model did not significantly differ from ordinary least squares regression, thus reflecting stability of the association reported herein. The robust model treated data points outside the confidence interval as outliers, resulting in particularly low weights for two data points. Furthermore, weights of data points were iteratively and individually adjusted according to the bisquare weighting function of the model. Fig. 3 contrasts both models, i.e. least squares regression and robust regression.

Table 2
Pearson's correlations (one-tailed) between time and frequency domain HRV parameters and BCI-performance (correctly copy-spelled letters [%]). Data of LF/HF were normalized using natural logarithm to achieve a normal distribution. For all other parameters data were of normal distribution and are thus reported on non-normalized data.

	N	r	p
RMSSD	34	.238	.087
pNN50	34	.235	.091
LF Power [%]	34	−.058	.373
HF Power [%]	34	.400	.01
LF/HFln	34	−.428	.006
RSAnorm	34	.506	.001

4. Discussion

The present study was designed to investigate the association between performance in the P300-BCI and resting HRV as an indicator of inhibitory control. The results demonstrated that vagally mediated HRV is positively associated with performance in the P300-BCI.

The association was particularly marked for RSAnorm. This normalized high frequency HRV parameter has been found to be much less influenced by sympathetic activity than without normalization (for a review Grossman and Taylor, 2007). When stimulating vagal and sympathetic pathways, Hedman et al. (1995) found RSAnorm reduced by factor 1.8 as compared to vagal stimulation only which confirmed remaining sympathetic influence on HF-HRV. Nevertheless, their results showed that without normalization, sympathetic influence reduced HF-HRV by factor 11.8 (see also Grossman and Taylor, 2007), i.e. normalizing HF-HRV (RSAnorm) reduces sympathetic influence by factor 6.6.

The remaining sympathetic effects may be the reason why correlations of performance with parameters of the time domain did not reach significance. This speculation is supported by the weak association between performance and LF-HRV, a parameter known to be affected by sympathetic activity (e.g. Pomeranz et al., 1985; Thayer et al., 2010, 2011). There is ongoing debate about the degree to which LF-HRV reflects sympathetic influence; some authors, for example, have reported far more substantial vagal influences or even doubt sympathetic presence in LF-HRV altogether (e.g. Eckberg, 1997; Porges, 2007). Our data indirectly support remaining sympathetic presence in LF-HRV.

The reported results in the association of vagally mediated HRV and BCI performance are not affected by sex and age differences.

Results from the robust regression model indicate that the predictive power of HRV for BCI performance is robust against outliers. Individual differences in HRV have been previously shown to be associated with cognitive performance requiring executive function (Thayer et al., 2009). All of these studies used working memory tests or Go-NoGo paradigms or both as cognitive tasks, which typically require overt motor responses to target stimuli presented on a screen, e.g., pressing a key on a keyboard. The present results are in line with these findings, confirming the notion of vagally mediated HRV to be crucial for executive function. Nevertheless, to our knowledge the present study is the first to involve a solely mental task not requiring overt motor responses to target stimuli, but a more direct, psychophysiological measure of inhibitory function, i.e. the P300-BCI. The present results, therefore, extend previous findings and provide even more convincing evidence of HRV's capacity to index prefrontal inhibition in tasks requiring executive control.

Beyond the theoretical implications for models of neurovisceral integration (Thayer and Lane, 2009) and executive function (e.g., Funahashi, 2001), the current results may become important for the use of BCI in clinical settings, provided they can be replicated in patient samples. At present, BCI is often targeted for people for whom conventional assistive communication technologies are not effective, because severe motor disabilities preclude their use of voluntary muscular control which depend on conventional methods of assistive technology (Kübler and Neumann, 2005). These include, for example, patients with amyotrophic lateral sclerosis (ALS), who decide to accept artificial ventilation and may thus render the locked-in state in which only residual muscular movement is preserved, individuals with severe cerebral palsy, brainstem stroke, or severe muscular dystrophies (Kübler and Birbaumer, 2008). On the one hand, the effect of BCI technologies on such severely paralyzed patients' quality of life depends on further improvements in the ease and convenience of their daily use (Zickler et al., 2009, 2011). On the other hand, interindividual differences also affect successful use of BCIs. Given its ease of application, it is conceivable to employ HRV as a predictive tool to assess individuals' capacity to use the P300-BCI successfully. An accurate assessment of expected individual BCI performance that requires a single monitoring of heart rate over a 5-minute period would aid differential indication and simplify BCI-based interventions designed to improve patients' quality of life. Nevertheless, the current data were obtained from a healthy sample; whether the results can be extended to a clinical population remains to be shown in future research.

Taken together, resting HRV was identified as a significant predictor of performance in the P300-BCI, which explained about 26% of the variance. Future studies should investigate whether resting HRV can also predict performance in patients with ALS or otherwise severely motor impaired patients and age matched healthy controls. Identifying predictors of BCI performance is still in its infancy and researchers mainly addressed performance prediction in a BCI based on sensorimotor rhythms that are regulated by motor imagery (Hammer et al., 2011; Kübler et al., 2011). Recently, Blankertz and colleagues identified the amplitude of resting state sensorimotor rhythms of the EEG as predictor of performance in motor imagery based BCI (Blankertz et al., 2010). Furthermore, Halder and colleagues were able to show that performance in the same motor imagery based BCI correlated strongly ($r = 0.7$) with activation in the dorsolateral prefrontal cortex, an area known to be involved in tasks requiring executive function (Halder et al., 2011). Kübler and colleagues found that the time needed to achieve significant performance with a BCI controlled by means of regulation of slow cortical potentials of the brain moderately predicted later performance above 70% correct responses which is considered the minimum accuracy for satisfying communication (Kübler et al., 2001b, 2004). However, predictors of performance in the P300 based BCI remain to be established. The present study is an attempt to fill this gap and to introduce HRV, a proxy for prefrontal cortical inhibition and executive function, as a predictor for P300-BCI performance.

Acknowledgments

Supported by the Deutsche Forschungsgemeinschaft (DFG; KU 1453/3-1) and the European ICT Programme Project FP7-224631. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

References

Ahern, G.L., Sollers, J.J., Lane, R.D., Labiner, D.M., Herring, A.M., Weinand, M.E., Hutzler, R., Thayer, J.F., 2001. Heart rate and heart rate variability changes in the intracarotid sodium amobarbital test. Epilepsia 42 (7), 912–921.

Allen, J.J.B., Chambers, A.S., Towers, D.N., 2007. The many metrics of cardiac chronotropy: a pragmatic primer and a brief comparison of metrics. Biological Psychology 74 (2), 243–262. doi:10.1016/j.biopsycho.2006.08.005.

Appelhans, B.M., Luecken, L.J., 2006. Heart rate variability as an index of regulated emotional responding. Review of General Psychology 10 (3), 229–240. doi:10.1037/1089-2680.10.3.229.

Appelhans, Bradley M., Luecken, L.J., 2008. Heart rate variability and pain: associations of two interrelated homeostatic processes. Biological Psychology 77 (2), 174–182. doi:10.1016/j.biopsycho.2007.10.004.

Benarroch, E.E., 1993. The central autonomic network: functional organization, dysfunction, and perspective. Mayo Clinic Proceedings. Mayo Clinic 68 (10), 988–1001.

Benarroch, E.E., 1997. The central autonomic network, Clinical Autonomic Disorders, 2nd ed. , pp. 17–23.

Berntson, G.G., Stowell, J.R., 1998. ECG artifacts and heart period variability: don't miss a beat! Psychophysiology 35 (1), 127–132.

Berntson, G.G., Quigley, K.S., Jang, J.F., Boysen, S.T., 1990. An approach to artifact identification: application to heart period data. Psychophysiology 27 (5), 586–598.

Birbaumer, N., Cohen, L.G., 2007. Brain-computer interfaces: communication and restoration of movement in paralysis. The Journal of Physiology 579 (Pt 3), 621–636. doi:10.1113/jphysiol.2006.125633.

Blankertz, B., Sannelli, C., Halder, S., Hammer, E.M., Kübler, A., Müller, K.-R., Curio, G., Dickhaus, T., 2010. Neurophysiological predictor of SMR-based BCI performance. NeuroImage 51 (4), 1303–1309. doi:10.1016/j.neuroimage.2010.03.022.

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 (5), 056013. doi:10.1088/1741-2560/7/5/056013.

Donchin, E., 1981. Surprise!…Surprise? Psychophysiology 18 (5), 493–513.

Eckberg, D.L., 1997. Sympathovagal balance: a critical appraisal. Circulation 96 (9), 3224–3232.

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 (6), 510–523.

Friedman, B.H., Thayer, J.F., 1998. Autonomic balance revisited: panic anxiety and heart rate variability. Journal of Psychosomatic Research 44 (1), 133–151.

Funahashi, S., 2001. Neuronal mechanisms of executive control by the prefrontal cortex. Neuroscience Research 39 (2), 147–165.

Grossman, P., Taylor, E.W., 2007. Toward understanding respiratory sinus arrhythmia: relations to cardiac vagal tone, evolution and biobehavioral functions. Biological Psychology 74 (2), 263–285. doi:10.1016/j.biopsycho.2005.11.014.

Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., Gramatica, F., Edlinger, G., 2009. How many people are able to control a P300-based brain-computer interface (BCI)? Neuroscience Letters 462 (1), 94–98. doi:10.1016/j.neulet.2009.06.045.

Halder, S., Agorastos, D., Veit, R., Hammer, E.M., Lee, S., Varkuti, B., Bogdan, M., Rosenstiel, W., Birbaumer, N., Kübler, A., 2011. Neural mechanisms of brain-computer interface control. NeuroImage 55 (4), 1779–1790. doi:10.1016/j.neuroimage.2011.01.021.

Halgren, E., Squires, N.K., Wilson, C.L., Rohrbaugh, J.W., Babb, T.L., Crandall, P.H., 1980. Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. Science (New York, N.Y.) 210 (4471), 803–805.

Hammer, Eva M., Halder, S., Blankertz, B., Sannelli, C., Dickhaus, T., Kleih, S., Müller, K.-R., Kübler, A., 2011. Psychological predictors of SMR-BCI performance. Biological Psychology. doi:10.1016/j.biopsycho.2011.09.006.

Hansen, A.L., Johnsen, B.J., Thayer, J.F., 2003. Vagal influence on working memory and attention. International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology 48 (3), 263–274.

Hayano, J., Sakakibara, Y., Yamada, M., Ohte, N., Fujinami, T., Yokoyama, K., Watanabe, Y., Takata, K., 1990. Decreased magnitude of heart rate spectral components in coronary artery disease. Its relation to angiographic severity. Circulation 81 (4), 1217–1224.

Hedman, A.E., Tahvanainen, K.U., Hartikainen, J.E., Hakumäki, M.O., 1995. Effect of sympathetic modulation and sympatho-vagal interaction on heart rate variability in anaesthetized dogs. Acta Physiologica Scandinavica 155 (2), 205–214.

Hoffmann, U., Vesin, J.-M., Ebrahimi, T., Diserens, K., 2008. An efficient P300-based brain-computer interface for disabled subjects. Journal of Neuroscience Methods 167 (1), 115–125. doi:10.1016/j.jneumeth.2007.03.005.

Kaufmann, T., Sütterlin, S., Schulz, S.M., Vögele, C., 2011. ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. Behavior Research Methods 43 (4), 1161–1170. doi:10.3758/s13428-011-0107-7.

Kleih, S.C., Kaufmann, T., Zickler, C., Halder, S., Leotta, F., Cincotti, F., Aloise, F., Riccio, A., Herbert, C., Mattia, D., Kübler, A., 2011. Out of the frying pan into the fire—the P300-based BCI faces real-world challenges. Progress in Brain Research 194, 27–46. doi:10.1016/B978-0-444-53815-4.00019-4.

Kübler, A., Birbaumer, N., 2008. Brain-computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology 119 (11), 2658–2666. doi:10.1016/j.clinph.2008.06.019.

Kübler, A., Neumann, N., 2005. Brain-computer interfaces—the key for the conscious brain locked into a paralyzed body. Progress in Brain Research 150, 513–525. doi:10.1016/S0079-6123(05)50035-9.

Kübler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J.R., Birbaumer, N., 2001a. Brain-computer communication: unlocking the locked in. Psychological Bulletin 127 (3), 358–375.

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 (11), 1533–1539.

Kübler, A., Neumann, N., Wilhelm, B., Hinterberger, T., Birbaumer, N., 2004. Predictability of brain-computer communication. Journal of Psychophysiology 18 (2–3), 121–129.

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. doi:10.1111/j.1749-6632.2008.04122.x.

Kübler, A., Blankertz, B., Müller, K.-R., Neuper, C., 2011. A model of BCI-control. Proceedings of the 5th International Brain-Computer Interface Conference 2011, pp. 100–103.

Mak, J.N., Arbel, Y., Minett, J.W., McCane, L.M., Yuksel, B., Ryan, D., Thompson, D., Bianchi, L., Erdogmus, D., 2011. Optimizing the P300-based brain-computer interface: current status, limitations and future directions. Journal of Neural Engineering 8 (2), 025003. doi:10.1088/1741-2560/8/2/025003.

Nijboer, F., Sellers, E.W., Mellinger, J., Jordan, M.A., Matuz, T., Furdea, A., Halder, S., Mochty, U., Krusienski, D.J., Vaughan, T.M., Wolpaw, J.R., Birbaumer, N., Kübler, A., 2008. A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology 119 (8), 1909–1916. doi:10.1016/j.clinph.2008.03.034.

Niskanen, J.-P., Tarvainen, M.P., Ranta-Aho, P.O., Karjalainen, P.A., 2004. Software for advanced HRV analysis. Computer Methods and Programs in Biomedicine 76 (1), 73–81. doi:10.1016/j.cmpb.2004.03.004.

Picton, T.W., 1992. The P300 wave of the human event-related potential. Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society 9 (4), 456–479.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology 118 (10), 2128–2148. doi:10.1016/j.clinph.2007.04.019.

Polich, J., Alexander, J.E., Bauer, L.O., Kuperman, S., Morzorati, S., O'Connor, S.J., Porjesz, B., Rohrbaugh, J., Begleiter, H., 1997. P300 topography of amplitude/latency correlations. Brain Topography 9 (4), 275–282.

Pomeranz, B., Macaulay, R.J., Caudill, M.A., Kutz, I., Adam, D., Gordon, D., Kilborn, K.M., Barger, A.C., Shannon, D.C., Cohen, R.J., 1985. Assessment of autonomic function in

humans by heart rate spectral analysis. The American Journal of Physiology 248 (1 Pt 2), H151–H153.
Porges, S.W., 2007. The polyvagal perspective. Biological Psychology 74 (2), 116–143. doi:10.1016/j.biopsycho.2006.06.009.
Saul, J., 1990. Beat-to-beat variations of heart rate reflect modulation of cardiac autonomic outflow. Physiology 5 (1), 32–37.
Schalk, Gerwin, McFarland, D.J., Hinterberger, T., Birbaumer, N., Wolpaw, J.R., 2004. BCI2000: a general-purpose brain-computer interface (BCI) system. IEEE Transactions on Bio-Medical Engineering 51 (6), 1034–1043. doi:10.1109/TBME.2004.827072.
Sellers, E.W., Donchin, E., 2006. A P300-based brain-computer interface: initial tests by ALS patients. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology 117 (3), 538–548. doi:10.1016/j.clinph.2005.06.027.
Sellers, E.W., Vaughan, T.M., Wolpaw, J.R., 2010. A brain-computer interface for long-term independent home use. Amyotrophic Lateral Sclerosis: Official Publication of the World Federation of Neurology Research Group on Motor Neuron Diseases 11 (5), 449–455. doi:10.3109/17482961003777470.
Silvoni, S., Volpato, C., Cavinato, M., Marchetti, M., Priftis, K., Merico, A., Tonin, P., Koutsikos, K., Beverina, F., Piccione, F., 2009. P300-based brain-computer interface communication: evaluation and follow-up in amyotrophic lateral sclerosis. Frontiers in Neuroscience 3, 60. doi:10.3389/neuro.20.001.2009.
Steiger, J., 1980. Tests for comparing elements of a correlation matrix. Psychological Bulletin 87 (2), 245–251. doi:10.1037/0033-2909.87.2.245.
Sütterlin, S., Herbert, C., Schmitt, M., Kübler, A., Vögele, C., 2011a. Frames, decisions, and cardiac-autonomic control. Social Neuroscience 6 (2), 169–177. doi:10.1080/17470919.2010.495883.
Sütterlin, S., Herbert, C., Schmitt, M., Kübler, A., Vögele, C., 2011b. Overcoming selfishness: reciprocity, inhibition, and cardiac-autonomic control in the ultimatum game. Frontiers in Psychology 2, 173. doi:10.3389/fpsyg.2011.00173.
Sutton, S., Braren, M., Zubin, J., John, E.R., 1965. Evoked-potential correlates of stimulus uncertainty. Science (New York, N.Y.) 150 (700), 1187–1188.
Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. European Heart Journal 17 (3), 354–381.
Thayer, J.F., Brosschot, J.F., 2005. Psychosomatics and psychopathology: looking up and down from the brain. Psychoneuroendocrinology 30 (10), 1050–1058. doi:10.1016/j.psyneuen.2005.04.014.
Thayer, J.F., Lane, R.D., 2000. A model of neurovisceral integration in emotion regulation and dysregulation. Journal of Affective Disorders 61 (3), 201–216.
Thayer, J.F., Lane, R.D., 2009. Claude Bernard and the heart-brain connection: further elaboration of a model of neurovisceral integration. Neuroscience and Biobehavioral Reviews 33 (2), 81–88. doi:10.1016/j.neubiorev.2008.08.004.
Thayer, J.F., Siegle, G.J., 2002. Neurovisceral integration in cardiac and emotional regulation. IEEE Engineering in Medicine and Biology Magazine 21 (4), 24–29. doi:10.1109/MEMB.2002.1032635.
Thayer, J.F., Hansen, A.L., Saus-Rose, E., Johnsen, B.H., 2009. Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine 37 (2), 141–153. doi:10.1007/s12160-009-9101-z.
Thayer, J.F., Yamamoto, S.S., Brosschot, J.F., 2010. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. International Journal of Cardiology 141 (2), 122–131. doi:10.1016/j.ijcard.2009.09.543.
Thayer, J.F., Loerbroks, A., Sternberg, E.M., 2011. Inflammation and cardiorespiratory control: the role of the vagus nerve. Respiratory Physiology & Neurobiology 178 (3), 387–394. doi:10.1016/j.resp. 2011.05.016.
Treder, M.S., Blankertz, B., 2010. (C)overt attention and visual speller design in an ERP-based brain-computer interface. Behavioral and Brain Functions: BBF 6, 28. doi:10.1186/1744-9081-6-28.
Valkonen-Korhonen, M., Tarvainen, M.P., Ranta-Aho, P., Karjalainen, P.A., Partanen, J., Karhu, J., Lehtonen, J., 2003. Heart rate variability in acute psychosis. Psychophysiology 40 (5), 716–726. doi:10.1111/1469-8986.00072.
Verleger, R., 1988. Event-related brain potentials and memory. The Behavioral and Brain Sciences 11, 343–427.
Zickler, C., Di Donna, V., Kaiser, V., Al-Khodairy, A., Kleih, S., Kübler, A., 2009. BCI applications for people with disabilities: defining user needs and user requirements. In: Emiliani, P.L., Burzagli, L., Como, A., Gabbanini, F., Salminen, A.-L. (Eds.), Assistive Technology from Adapted Equipment to Inclusive Environments, AAATE. : Assistive Technology Research Series, 25. IOS Press, Amsterdam, pp. 185–189.
Zickler, C., Riccio, A., Leotta, F., Hillian-Tress, S., Halder, S., Holz, E., Staiger-Sälzer, P., Hoogerwerft, 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 24 (4), 222.

2.2 Kaufmann, T., Sütterlin, S., Schulz, S. M., and Vögele, C. (2011). ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. Behavior Research Methods

Abstract: The importance of appropriate handling of artifacts in interbeat interval (IBI) data must not be underestimated. Even a single artifact may cause unreliable heart rate variability (HRV) results. Thus, a robust artifact detection algorithm and the option for manual intervention by the researcher form key components for confident HRV analysis. Here, we present ARTiiFACT, a software tool for processing electrocardiogram and IBI data. Both automated and manual artifact detection and correction are available in a graphical user interface. In addition, ARTiiFACT includes time- and frequency-based HRV analyses and descriptive statistics, thus offering the basic tools for HRV analysis. Notably, all program steps can be executed separately and allow for data export, thus offering high flexibility and interoperability with a whole range of applications.

Link to publication source: Behavior Research Methods, Springer Science+Business Media B.V.

Behav Res (2011) 43:1161–1170
DOI 10.3758/s13428-011-0107-7

ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis

Tobias Kaufmann · Stefan Sütterlin · Stefan M. Schulz · Claus Vögele

Published online: 14 May 2011

Abstract The importance of appropriate handling of artifacts in interbeat interval (IBI) data must not be underestimated. Even a single artifact may cause unreliable heart rate variability (HRV) results. Thus, a robust artifact detection algorithm and the option for manual intervention by the researcher form key components for confident HRV analysis. Here, we present ARTiiFACT, a software tool for processing electrocardiogram and IBI data. Both automated and manual artifact detection and correction are available in a graphical user interface. In addition, ARTiiFACT includes time- and frequency-based HRV analyses and descriptive statistics, thus offering the basic tools for HRV analysis. Notably, all program steps can be executed separately and allow for data export, thus offering high flexibility and interoperability with a whole range of applications.

Keywords Heart rate variability · ECG artifact processing · Electrocardiogram · Analysis software

In recent years, heart rate variability (HRV) has been increasingly used in medical and psychological research as a noninvasive method for the reliable estimation of vagally mediated modulation of sinus node cardiac activity. On the one hand, reduced vagal-cardiac control (i.e., low HRV) has been demonstrated to be a major risk factor for various conditions, particularly cardiovascular disorders (Kamath et al., 1987). On the other hand, high HRV has been shown to be a valid indicator of prefrontal inhibitory capacity involved in improved executive functions, such as attentional processing or impulse control (Appelhans & Luecken, 2006; Thayer & Brosschot, 2005), and in a variety of psychopathological conditions with symptoms of impaired behavioral and emotional regulation (Thayer & Lane, 2009; Schulz, Alpers, & Hofmann, 2008). The reciprocal interconnections of the central autonomic network (Bennarroch, 1993) further highlight the clinical relevance of HRV due to a rising number of HRV biofeedback studies (for a review, see Wheat & Larkin, 2010) and techniques available to health care institutions.

There are expensive commercial HRV analysis programs such as Nevrokard® (Nevrokard, Izola/Slovenia), and powerful freely available analysis tools, such as Kubios HRV (University of Eastern Finland, Kuopio/Finland; Niskanen, Tarvainen, Ranta-aho, & Karjalainen, 2004). Nevertheless, standardization and unification of data processing have not yet resulted in the development of freely available stand-alone software comprising all necessary steps, from processing of the electrocardiogram (ECG) to IBI, detection and correction of artifacts, as well as HRV and statistical data analysis. In the present article, we describe a program that includes all of these steps. A particular difference in comparison with existing software solutions is an implemented artifact detection algorithm that is based on the intra-individual calculation of a threshold criterion and its transparent treatment.

T. Kaufmann (✉) · S. M. Schulz
Department of Psychology I, University of Würzburg,
Marcusstr. 9–11,
97070 Würzburg, Germany
e-mail: tobias.kaufmann@uni-wuerzburg.de

S. Sütterlin · C. Vögele
Research Unit INSIDE, University of Luxembourg,
Campus Walferdange, Route de Diekirch - BP2,
7220 Walferdange, Luxembourg

Artifact handling

Artifact detection Assessment of HRV requires the extraction of IBI, which is usually performed by extracting R-peaks from digitized ECG data. In cases in which there is no raw ECG available (e.g., when recording devices such as the POLAR® RS800 [POLAR, Kempele/Finland] provide already extracted IBIs; Nunan et al., 2009), correct detection and correction of artifacts poses a particular challenge. Movement artifacts, technical failure, and poor data quality—particularly during ambulatory monitoring—increase the need for sophisticated artifact processing. Importantly, the potentially fatal impact of even minor contamination of data with artifacts (e.g., an ectopic beat resulting in a markedly shorter IBI) cannot be stressed enough (Berntson & Stowell, 1998). Despite the high importance of the particular method used for artifact detection, these procedures are rarely reported in studies assessing HRV data. The common practice of only visual inspection of extra or missing IBIs is highly questionable, in terms of reliability and practicality. Furthermore, the use of threshold values is questionable, since this gives rise to either Type I, or, more likely, Type II errors, especially when unaccounted interindividual differences in baseline heart rate (HR) are present. We argue, therefore, that the choice of artifact processing methods is often poorly justified in the literature. Importantly, such lack of consistency impairs replication and increases the risk of arbitrary decisions, particularly by inexperienced novices following these procedures.

Berntson, Quigley, Jang, and Boysen (1990) developed an artifact detection algorithm based on individual threshold criteria of artificial beats, derived from individual IBI distributions and their estimated real (not contaminated) distribution. This procedure allows for a highly automated and reliable data inspection with large data sets. Although this algorithm can still be considered as the unrivaled state-of-the-art method for IBI-based artifact detection, to our knowledge, it has not been implemented in freely available data processing software. This algorithm would be especially useful in a freely available, stand-alone software program covering the whole range of data processing from raw ECG to HRV parameters.

Artifact correction Several approaches for correcting artifacts in IBI series have been suggested, including linear and cubic spline interpolation, nonlinear predictive interpolation, and exclusion of ectopy-containing data segments (Lippman, Stein, & Lerman, 1994). Although promising alternatives based on artifact original IBI data alone have been reported (e.g., Clifford & Tarassenko, 2005), linear and cubic spline interpolation are still considered standard procedures for the replacement of missing or incorrect IBI, as was suggested by Malik and Camm (1995). Kubios HRV applies cubic spline interpolation to replace missing IBIs. Therefore, running comparisons against this software should rely on the same interpolation method. It appears, therefore, appropriate to include different options for artifact treatment available within one software package.

ARTiiFACT provides a powerful tool for processing of ECG and IBI data and calculation of all basic HRV parameters. In contrast with other freely available software solutions, ARTiiFACT uses independent modules allowing for various import and export opportunities and the usage of single modules only. Artifact detection is based on an automatic and intra-individually applied algorithm with excellent performance, as compared with widely used alternatives.

Computational methods and theory

ARTiiFACT has been designed to provide researchers with a software tool covering the complete range of data processing steps, from raw ECG data to deriving HRV parameters for statistical analysis. ARTiiFACT provides (a) import options for raw ECG data as well as IBI data, (b) automated artifact detection based on distribution-related criteria determined for each individual data set, (c) interpolation of missing data (optional choice of linear or cubic spline interpolation or deletion of artifact IBI), (d) calculation of common HRV parameters in both time and frequency domains, and (e) exports for all steps of data processing via spreadsheets or text files. Notably, ARTiiFACT offers a convenient data interface to RSAtoolbox (Schulz, Ayala, Dahme, & Ritz, 2009), a freely available implementation of peak-valley analysis of RSA, including a regression-based correction of within-individual effects of breathing on estimates of vagal tone, given that the appropriate respiratory parameters (i.e., duration and volume per breath) are available (on respiratory control see e.g., Grossman & Taylor, 2007, Ritz & Dahme 2006). It is also possible to estimate breathing parameters directly from the IBI time series (e.g., O'Brien & Heneghan, 2007). This allows for partial control of respiratory effects on HRV without even recording breathing. The various I/O options of ARTiiFACT make it particularly easy to apply such corrections before submitting data to the HRV analysis module. Particular emphasis was put on intuitive handling, lean program structure, and manual intervention options. The software is MATLAB® based and is available as a stand-alone 32-bit Windows® application.

R-peak detection ARTiiFACT provides the possibility to low-pass or high-pass filter raw ECG data at a manually adjustable cut-off frequency. Furthermore, a window-based linear detrending method was implemented in order to purge data from long time drifts. For R-peak detection, there is a choice between global or local threshold detection criteria (see Fig. 1). Using global threshold detection, the software suggests minimum peak amplitude based on available ECG data that may be adjusted manually by the user. R-peaks exceeding this amplitude criterion are

Behav Res (2011) 43:1161–1170

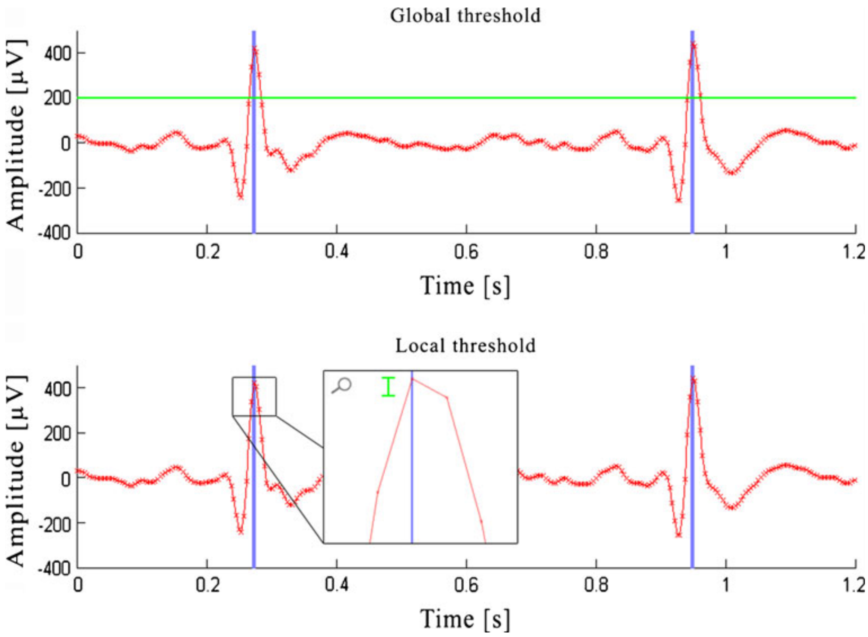


Fig. 1 Illustration of global and local threshold detection

detected throughout the data set. However, when drifts are present in the data, the user may switch to local threshold detection. Local thresholds are defined as the minimum voltage difference between two neighboring data points at a peak. Together, with a predefined minimum R–R distance, this allows for robust R-peak detection. If a data point exceeds the amplitude of its two surrounding points by more than the threshold criterion, and if the preceding R-peak has a greater distance than the minimum R-R distance, this value is identified as an R-peak.

Artifact detection and treatment The identification of spurious IBI—for example, that resulting from erroneous raw data, movement artifacts, or equipment failure—is implemented using the artifact detection algorithm proposed by Berntson et al. (1990). In the literature, manual handling of artifacts is typically based on rather subjective criteria, or, similarly, on arbitrary threshold definitions in which the artifact criteria are detected when exceeding a uniformly applied threshold (e.g., 130% above or below mean/median). Because of varying HR, both methods are not applicable across individuals and are inherently biased by existing artifacts, since they are based on data distributions, including artifacts. ARTiiFACT derives the artifact detection criterion from the distribution of IBI differences of the individual subject and applies percentile-based distribution indices, which are less sensitive to corruption by the presence of artifacts than simple threshold criteria or least-square estimates. Next, ARTiiFACT assesses the distribution parameters of an individual's IBI dataset, removes artifacts in the first and fourth quartile, and estimates the overall (artifact-free) standard deviation on the basis of the interquartile range. On the basis of these data, the final calculation of an individual threshold criterion for beat-to-beat differences to identify artifacts is completed. Artifacts can be treated in two ways: either (a) deletion or (b) estimation, which is often done by linear or cubic spline interpolation (see Lippman et al., 1994, for a comparison of common approaches). Deleting artifacts prevents incorrect estimation of artifact IBIs but inevitably crops the data set. This reduces data reliability and may bias the data, especially when artifacts are correlated systematically with experimental conditions. In contrast, interpolation maintains both, the length and structural characteristics of the IBI series, but bears the risk of misestimating the inserted IBIs. Therefore, which of the different estimate techniques performs best may depend on the particular application.

Time-domain methods Commonly used time-domain parameters, such as SDNN, RMSSD, NN50, pNN50 (for a definition, see Allen, Chamber, & Tower, 2007) are calculated from the IBI time series (see Table 1).

Frequency domain methods Spectral frequency measures are derived using the Fast Fourier Transformation (FFT). Frequency bands are delimited in line with the Task Force's (1996) recommendations as high frequency (HF, 0.15–0.4 Hz), low frequency (LF, 0.04–0.15 Hz), and very low frequency (VLF, < .04 Hz). These frequency band definitions are default values and can be altered manually.

Table 1 Time- and frequency-domain measures

hrvAnalysis	
Time domain	
Mean RR	Mean of interbeat intervals
Median RR	Median of interbeat intervals
SDNN	Standard deviation of interbeat intervals
RMSSD	Root mean square of successive interbeat intervals
NN50	Number of interbeat intervals that differ in more than 50 ms
pNN50	NN50 expressed as a percentage
Frequency domain	
VLF [ms²]	Very-low-frequency component
LF [ms²]	Low-frequency component
HF [ms²]	High-frequency component
LF [n.u.]	Normalized units of LF
HF [n.u.]	Normalized units of HF
LF/HF	Ratio of low and high frequency

As a default, the FFT applies a Hanning window of a 256-s width, an interpolation rate of 4 Hz (spline interpolation), and an overlap of 50% to the resampled and detrended data (method of least squares). All FFT parameters can be altered manually.

Measures of dispersion To allow for visual evaluation of data quality and assumptions required for further data analysis (e.g., subsequent multivariate testing), standard measures of dispersion such as standard deviation, variance, and range are displayed, as well as parameters of skewness and kurtosis. The IBI distribution is shown in a histogram. Mean absolute deviation is given as a measure of dispersion less sensitive to outliers. Furthermore, the interquartile range, defined as the range between second and third quartile (Q3–Q2), and the Kolmogorov–Smirnov test for normality of the IBI distribution (alpha set to .05) are provided.

Program description

ARTiiFACT consists of four subcomponents that are all callable from the main application window.

1. ecgExtract: Extraction of IBI from continuous ECG recordings.
2. ibiArtifactProcessing: Artifact detection and correction in IBI data.
3. hrvAnalysis: Calculation of HRV in time and frequency domains; see Table 1 for a list of parameters.
4. distributionStatistics: Tests for normality and measures of dispersion for the IBI distribution.

The ARTiiFACT software thus covers the full range of steps from processing raw ECG via IBI extraction to data cleaning and statistical analysis. Moreover, ARTiiFACT enables users to exploit output from intermediate steps separately. Thus, all subcomponents can be executed independently, and their input requirements and output do not depend on or interfere with each other. For example, processing of artifacts in IBI data can be performed without any further analysis, and in this case, ARTiiFACT allows exporting artifact-corrected IBIs to a file without further processing. This program structure of independent, but seamlessly interlinked, subcomponents serves the principle of transparency and flexibility. Alternatively, users may perform a complete data analysis with ARTiiFACT using its graphical user interface (GUI). The didactic value of the structured approach to HRV analysis with this interface may establish ARTiiFACT as a reliable and transparent tool for junior researchers and in teaching contexts, as well as in medical use.

The ecgExtract module The subcomponent ecgExtract (Fig. 2) allows for import of different file formats ("*.txt," "*.mat," "*.xls," "*.hdf5") containing data of the format (channels X samples). It is possible to select the channel number (i.e., column) containing ECG data, to skip any amount of header lines, and to individually set the sampling rate. The continuous ECG data are plotted alongside a scrollable axis. Data can optionally be cut, inverted, linearly detrended, and/or filtered (high pass/low pass). A threshold for the R-peak detection can be set manually, and the appropriate detection method can be selected. R-peaks are detected and plotted for visual inspection. In case of incorrectly detected peaks, manual intervention is possible.

The ibiArtifactProcessing module The subprogram ibiArtifactProcessing requires an input text file consisting of IBI data (one column, no header), that can optionally be cut to create a suitable epoch before further processing. Detection

Behav Res (2011) 43:1161–1170 1165

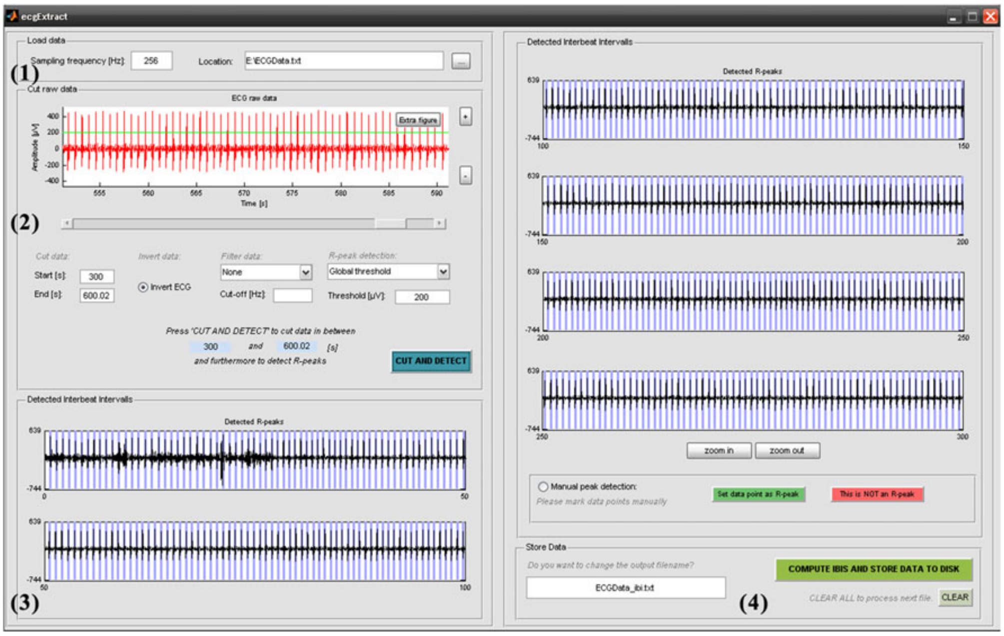


Fig. 2 The graphical user interface of the subprogram ecgExtract. 1 Definition of individual sampling rate. **2** Raw ECG data plot with several options (cut, filter, invert). **3** Plot of detected R-peaks with possibility of manual intervention. **4** Data storage

of artifacts is performed as described in the section Computational Methods. It is possible to check data visually in order to detect, revise, and deselect artifacts manually. For artifact correction, there is a choice between linear or cubic spline interpolation, as well as deletion of artifact data. Interpolation replaces the detected artifacts with estimates according to the chosen interpolation method. Deletion does not insert IBI estimates, thus cropping the IBI data series. Notably, for assessing artifact position, ARTiiFACT allows exporting a file with data flags indicating the position of artifacts. All steps are displayed in the GUI (Fig. 3) and may be altered retroactively until the desired end product is achieved and saved.

The hrvAnalysis module **HRV** analysis is performed in both the time and frequency domains, which provide several highly correlated parameters indicating the extent of HRV.

The dispersionStatistics module Individual HRV parameters are only as accurate as the experimental conditions and accuracy of data acquisition allow. Supervision of parameters of data quality and fidelity can provide information about the quality and accuracy of the statistical outcomes and, ultimately, the validity of their interpretation. This subprogram provides descriptive measures of dispersion such as skewness, kurtosis, interquartile range, and the Kolmogorov–Smirnov test for normality, as well as quantile–quantile-plots. These data can help to assess whether experimental conditions were appropriate and equipment was accurately used. Furthermore, it might help to identify sources of data corruption such as insufficient relaxation periods, nonstandardized stress conditions, or anxiety caused by the laboratory setting before and during data collection, which, for example, would result in a skewed distribution.

The subprograms hrvAnalysis and distributionStatistics provide the possibility to print a report sheet in the portable document format (Adobe® pdf, see Fig. 4).

Samples of typical program runs

ARTiiFACT was tested with data sets from both ECG (recorded with Ag–AgCl electrodes according to Einthoven lead I, at 256 Hz using a g.USBAmp [g.tec, Austria] amplifier) and IBI data (recorded with Polar RS800CX, Nunan et al., 2009). The following criteria were applied to these tests:

ECG raw data are sometimes contaminated with artifacts—for example, artifacts caused by movement. It appears that such artifacts are hardly distinguishable from valid R-peaks. An example (real participant data) is given in Fig. 5a, in which movement artifacts occur between seconds 8 and 10. This results in two similar peaks at t = 8.81 s and t = 8.88 s. In cases such as this,

Fig. 3 The graphical user interface of the subprogram ibiArtifactProcessing. **1** Load data. **2** Raw data plot on a scrollable axis and the possibility to cut data manually. **3** Data plot with detected artifacts and the possibility of manual intervention. **4** Data plot showing the corrected IBIs and the histogram. **5** Data storage and rsaToolbox export

visual inspection or completely automated ECG processing are not reliable procedures to detect and distinguish artifacts from correctly identified R-peaks. ARTiiFACT provides the possibility to deselect a detected peak

Fig. 4 The graphical user interface of hrvAnalysis and descriptiveStatistics. **1** Load data. **2** Insert participant ID (optional). **3** Raw data plot. **4** Settings for HRV computation. **5** Time domain analysis. **6** Frequency domain analysis. **7** Store data. **8** Distribution statistics

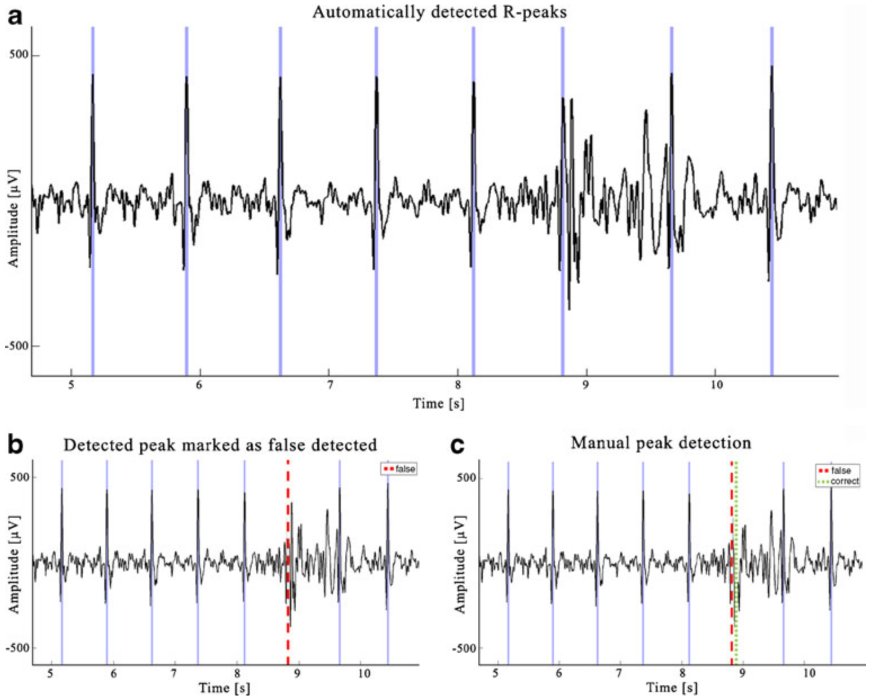


Fig. 5 Manual detection of R-peaks. **a** Raw data with automatically detected R-peaks. Between seconds 8 and 9, a wrong detection might have occurred, possibly due to movement artifacts in the data. ARTiiFACT gives the possibility of manual intervention, by **b** deletion of a detected R-peaks, and **c** insertion of R-peaks. However, if uncertain about the reliability of the detected R-peak, manual peak detection should not be performed, and ARTiiFACT's artifact correction should be used

(Fig. 5b). However, if, after inspection, a peak is to be defined as true, it can be manually defined as an R-peak (Fig. 5c). Thus, the software allows fully automated artifact detection as well as complete manual control, depending on the user's preference.

Computing HRV of IBI data crucially depends on artifact-free data sets (Berntson & Stowell, 1998), since any artifact would distort variability. The algorithm recommended by Berntson et al. (1990), therefore, tries to exclude any potential artifact with a sensitive algorithm before deriving the final criteria for identifying true artifacts. Figure 6 shows a set of IBI data containing three automatically detected artifacts. Kubios HRV applies cubic spline interpolation to replace missing IBIs. To allow for a straightforward comparison, cubic spline interpolation (blue line in Fig. 6) was used to replace artifacts when appropriate. As can be concluded from Fig. 6, the estimation fits precisely into the trend of interbeat variability.

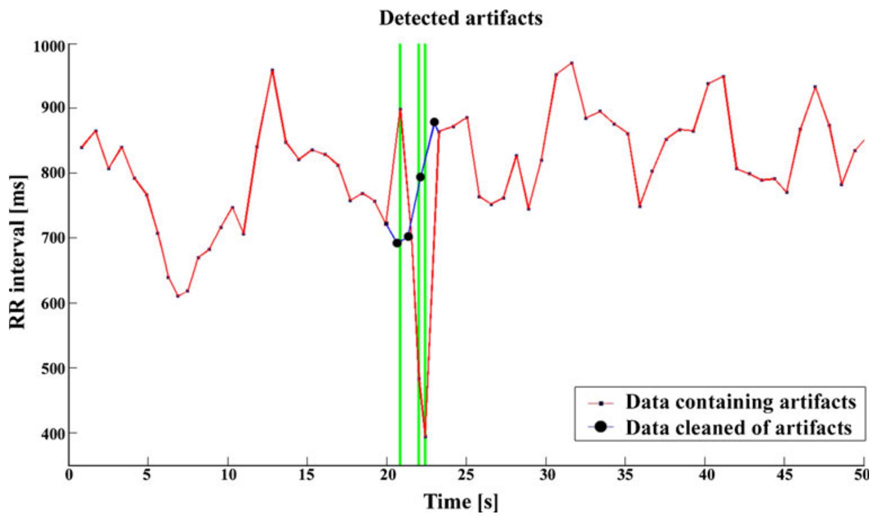


Fig. 6 Artifact detection and correction in IBI data. Between seconds 22 and 24, an artifact contaminates the raw data. ARTiiFACT detects all data points that are likely to be artifacts and replaces them by estimated IBIs (see text for details)

The automated artifact detection was further validated with artificial IBI data sets. These were created as follows: Starting at a mean IBI length of 833 ms, a random IBI was computed within the range of <100 ms and >5 ms in relation to the preceding IBI. This process was iterated for 100 times and resulted in a data set containing artificial IBIs. Moreover, two types of artifacts were included:

Type A: Two IBIs were combined to simulate a missed R-peak detection in calculating IBI data of raw ECG data. Type A IBIs are therefore about twice the length of a single typical IBI.
Type B: One IBI was split in half to simulate an artificial extra R-peak. Type B artifacts therefore result in two IBIs; their size can differ from equal length to major disparate lengths, which apparently is a challenge for artifact detection algorithms, since it may result in artificial intervals that are not distinguishable from valid intervals (Berntson & Stowell, 1998). For validation, different data sets were created in which the ratio of the two IBIs' sizes varied from 2.5–5% to 95–97.5%; 5–10% to 90–95%; 10–15% to 85–90%; 15–20% to 80–85%; and so on, up to 45–50% to 50–55% (the actual percentage value was chosen randomly within these constraints). This resulted in 10 different types of data sets.

Each type of data set was generated five times to avoid dependency of results on a single data set. Since each data set contained five artifacts of Type A and five artifacts of Type B, and since Type B means two artificial IBIs to detect per artifact type, 15 data points ought to have been detected in each data set.

Table 2 shows the result of this validation. In all cases, artifacts of Type A were reliably detected. For Type B artifacts, the shorter IBI was always correctly identified. As expected, the detection of the corresponding longer IBI depended on its size. IBIs larger than 85% of the original IBI were not distinguished from valid IBIs. However, for the next 10% decrease of IBI size, detection rate increased. Only a few false alarms occurred between 75 and 90%. If the larger IBI was smaller than 75% of the actual complete IBI, all IBIs were reliably detected as artifacts.

To validate the quality of this artifact detection algorithm, the same data sets were analyzed with the Kubios HRV software, a widely used tool that is often used as HRV analysis software (see, e.g., Culbertson et al., 2010; Sütterlin, Herbert, Schmitt, Kübler, & Vögele, 2011) and offers artifact correction at five levels of sensitivity ranging from *very low* to *very strong*. Table 3 summarizes the corresponding results for the three criteria: low, medium, and strong. Both low and medium correction criteria resulted in good detection of Type A artifacts, but only poor detection of Type B artifacts. Only the shorter part of Type B artifacts was reliably detected. Unfortunately, the undetected longer part still contaminated the remaining IBI series. Although the criterion strong resulted in better detection of both parts of Type B artifacts, it did not reach ARTiiFACT's detection accuracy (see Tables 2 and 3, number of missed detections). Moreover, with the criteria set to strong, the number of false detections increased dramatically throughout the whole test data set (44 false detections in total).

ARTiiFACT, by contrast, produced a considerably lower number of false detections (12 false detections in total). Their occurrence also was restricted to test data sets where the size of larger artificial IBIs was rather large (i.e., between 85 and 90%). In sum, ARTiiFACT showed a better artifact detection rate and fewer false detections.

Table 2 Validation of the automated artifact detection

Artifact Detection in ARTiiFACT						
Length of short IBI (% of actual IBI) for Type B artifacts	Number of Type A artifacts	Number of Type B artifacts	Number of false detections	Number of missed detections	*d'*	*ß*
2.5–5%	25/25	25/50	0	25	NA	NA
5–10%	25/25	25/50	0	25	NA	NA
10–15%	25/25	25/50	3	25	2,89	65,96
15–20%	25/25	31/50	2	19	3,32	90,69
20–25%	25/25	40/50	7	10	3,45	21,34
25–30%	25/25	50/50	0	0	NA	1
30–35%	25/25	50/50	0	0	NA	1
35–40%	25/25	50/50	0	0	NA	1
40–45%	25/25	50/50	0	0	NA	1

Different data sets were created, including artifacts with the whole range of possible artifact length. All data sets were validated according to the detection of two types of artifacts, either Type A or Type B artifacts. Both false alarms and true missings were validated. *d'* = sensitivity, *ß* = response bias

Table 3 Validation of artifact detection performance in Kubios HRV for comparison

Artifact Detection in Kubios HRV

Length of short IBI (% of actual IBI) for Type B artifacts	Number of Type A artifacts	Number of Type B artifacts	Number of false detections	Number of missed detections	*d'*	*ß*
LOW:						
2.5–5%	25/25	25/50	0	25	NA	NA
5–10%	25/25	25/50	0	25	NA	NA
10–15%	25/25	25/50	0	25	NA	NA
15–20%	25/25	25/50	0	25	NA	NA
20–25%	25/25	25/50	0	25	NA	NA
25–30%	25/25	25/50	0	25	NA	NA
30–35%	25/25	24/50	0	26	NA	NA
35–40%	25/25	26/50	0	24	NA	NA
40–45%	25/25	31/50	0	19		
MEDIUM:						
2.5–5%	25/25	25/50	0	25	NA	NA
5–10%	25/25	25/50	0	25	NA	NA
10–15%	25/25	25/50	0	25	NA	NA
15–20%	25/25	25/50	0	25	NA	NA
20–25%	25/25	25/50	0	25	NA	NA
25–30%	25/25	26/50	0	24	NA	NA
30–35%	24/25	33/50	0	17	NA	NA
35–40%	25/25	47/50	0	3	NA	NA
40–45%	25/25	50/50	0	0	NA	1
STRONG:						
2.5–5%	25/25	25/50	6	25	2,67	35,06
5–10%	25/25	25/50	2	25	3,02	95,74
10–15%	25/25	25/50	7	25	2,57	27,28
15–20%	25/25	27/50	4	23	2,90	50,34
20–25%	25/25	34/50	7	16	3,69	160,12
25–30%	25/25	47/50	1	3	4,09	7,34
30–35%	25/25	50/50	4	0	NA	0
35–40%	25/25	50/50	10	0	NA	0
40–45%	25/25	50/50	2	0	NA	0

Kubios HRV provides artifact detection at different criteria levels ranging from very low to very strong. Results presented for the three criteria: low, medium and strong. *d'* = sensitivity, *ß* = response bias

For further quality assessment of the artifact detection algorithm, we computed signal detection theory estimators for sensitivity (*d*') and detection bias (β) (Stanislaw & Todorov, 1999). Notably, ARTiiFACT reached the goal of β = 1 for those data sets, with the larger IBI being smaller than 75% of the actual complete IBI (see Table 2). Kubios HRV, on the other hand, reached β = 1 only for those data sets with larger IBI smaller than 60% of the actual IBI, and only for the detection criterion medium. For the detection criteria low and strong, it failed to reach the target β value (see Table 3).

Therefore, we conclude that ARTiiFACT provides sensitive, well-balanced, and reliable artifact detection in IBI data, thus optimizing the accuracy of subsequent analyses, such as computing HRV parameters (Berntson & Stowell, 1998).

Hardware and software specifications

ARTiiFACT was developed using MATLAB® 2009b and was compiled with the MATLAB® compiler 4.13. The necessary MATLAB® component runtime 7.13 (MCR) was packaged along with the application. ARTiiFACT should work on all 32-bit Windows operating systems (XP, Vista and 7 were tested). The minimal desktop resolution is 1280 × 768 pixels.

Mode of availability of program

ARTiiFACT is freely available upon request (email tobias.kaufmann@uni-wuerzburg.de). Additionally, the requesting user is provided with a tutorial and user manual. Registered users may opt in for notification of free updates. Current planning includes options for batch processing, extending HRV analysis by autoregressive models and nonlinear analysis, and a full module for treatment of respiration-related issues.

Summary

In the present article, we present a software tool providing the user with an efficient artifact detection algorithm, including the possibility of manual revision, artifact removal, and computation of HRV, as well as descriptive statistics on the distribution of data. Although a broad variety of settings and functions are available, ARTiiFACT provides a graphical user interface that makes it applicable for both research and teaching purposes. Its modular structure and compatibility allows integration with other software tools, replacing one or more of ARTiiFACT's subcomponents to maximize benefits by combining advantages of various software solutions.

Author Note The work presented in this paper was funded by a grant from the Alfried-Krupp-Science Foundation to C. V., grant number DM8-MEDFAK 01.

References

Allen, B., Chambers, A. S., & Towers, D. N. (2007). The many metrics of cardiac chronotropy: A pragmatic primer and a brief comparison of metrics. *Biological Psychology, 74,* 243–262.

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology, 10,* 229–240.

Bennarroch, E. E. (1993). The central autonomic network: Functional organization, dysfunction, and perspective. *Mayo Clinic Proceedings, 68,* 988–1001.

Berntson, G. G., Quigley, K. S., Jang, J. F., & Boysen, S. T. (1990). An approach to artifact identification: Application to heart period data. *Psychophysiology, 27,* 586–598.

Berntson, G. G., & Stowell, J. R. (1998). ECG artifacts and heart period variability: Don't miss a beat! *Psychophysiology, 35,* 125–132.

Clifford, G. D., & Tarassenko, L. (2005). Quantifying errors in spectral estimates of HRV due to beat replacement and resampling. *IEEE Transactions on Biomedical Engineering, 52,* 630–638.

Culbertson, C., Nicolas, S., Zaharovits, I., London, E. D., La Garza, I., Brody, A. L., et al. (2010). Methamphetamine craving induced in an online virtual reality environment. *Pharmacology, Biochemistry and Behavior, 96,* 454–460.

Grossman, P., & Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology, 74,* 263–285.

Kamath, M. V., Ghista, D. N., Fallen, E. L., Fitchett, D., Miller, D., & McKelvie, R. (1987). Heart rate variability power spectrogram as a potential noninvasive signature of cardiac regulatory system response, mechanisms, and disorders. *Heart and Vessels, 3,* 33–41.

Lippman, N., Stein, K. M., & Lerman, B. B. (1994). Comparison of methods for removal of ectopy in measurement of heart rate variability. *The American Journal of Physiology, 267,* H411–H418.

Malik, M., & Camm, A. J. (1995). *Heart rate variability*. New York, NY: Futura.

Niskanen, J. P., Tarvainen, M. P., Ranta-aho, P. O., & Karjalainen, P. A. (2004). Software for advanced HRV analysis. *Computer Methods and Programs in Biomedicine, 76,* 73–81.

Nunan, D., Donovan, G., Jakovljevic, D. G., Hodges, L., Sandercock, G. R. H., & Brodie, D. A. (2009). Validity and reliability of short-term heart-rate variability from the Polar S810. *Medicine and Science in Sports and Exercise, 41,* 243–250.

O'Brien, C., & Heneghan, C. (2007). A comparison of algorithms for estimation of a respiratory signal from the surface electrocardiogram. *Computers in Biology and Medicine, 37,* 305–314.

Ritz, T., & Dahme, B. (2006). Implementation and interpretation of respiratory sinus arrhythmia measures in psychosomatic medicine: Practice against better evidence? *Psychosomatic Medicine, 68,* 617–627.

Schulz, S. M., Alpers, G. W., & Hofmann, S. G. (2008). Negative self-focused cognitions mediate the effect of trait social anxiety on state anxiety. *Behavior Research and Therapy, 46,* 438–449.

Schulz, S. M., Ayala, E., Dahme, B., & Ritz, T. (2009). A MATLAB toolbox for correcting within-individual effects of respiration rate and tidal volume on respiratory sinus arrhythmia during variable breathing. *Behavior Research Methods, 41,* 1121–1126.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31,* 137–149.

Sütterlin, S., Herbert, C., Schmitt, M., Kübler, A., & Vögele, C. (2011). Frames, decisions, and cardiac–autonomic control. *Social Neuroscience, 6,* 169–177.

Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology. (1996). Heart rate variability—Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal, 17,* 354–381.

Thayer, J. F., & Brosschot, J. F. (2005). Psychosomatics and psychopathology: Looking up and down from the brain. *Psychoneuroendocrinology, 30,* 1050–1058.

Thayer, J. F., & Lane, R. D. (2009). Claude Bernard and the heart–brain connection: Further elaboration of a model of neurovisceral integration. *Neuroscience and Biobehavioral Reviews, 31,* 81–88.

Wheat, A. L., & Larkin, K. T. (2010). Biofeedback of heart rate variability and related physiology: A critical review. *Applied Psychophysiology and Biofeedback, 35,* 229–242.

2.3 Kaufmann, T., Völker S., Gunesch L. and Kübler A. (2012). Spelling is just a click away - a user-centered brain-computer interface including auto-calibration and predictive text entry. Frontiers in Neuroscience

Abstract: Brain-computer interfaces (BCI) based on event-related potentials (ERP) allow for selection of characters from a visually presented character-matrix and thus provide a communication channel for users with neurodegenerative disease. Although they have been topic of research for more than 20 years and were multiply proven to be a reliable communication method, BCIs are almost exclusively used in experimental settings, handled by qualified experts. This study investigates if ERP-BCIs can be handled independently by laymen without expert support, which is inevitable for establishing BCIs in end-user's daily life situations. Furthermore we compared the classic character-by-character text entry against a predictive text entry (PTE) that directly incorporates predictive text into the character-matrix. N = 19 BCI novices handled a user-centered ERP-BCI application on their own without expert support. The software individually adjusted classifier weights and control parameters in the background, invisible to the user (auto-calibration). All participants were able to operate the software on their own and to twice correctly spell a sentence with the auto-calibrated classifier (once with PTE, once without). Our PTE increased spelling speed and, importantly, did not reduce accuracy. In sum, this study demonstrates feasibility of auto-calibrating ERP-BCI use, independently by laymen and the strong benefit of integrating predictive text directly into the character-matrix.

Link to publication source: Frontiers in Neuroscience, Frontiers Media SA

frontiers in
NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 23 May 2012
doi: 10.3389/fnins.2012.00072

Spelling is just a click away – a user-centered brain–computer interface including auto-calibration and predictive text entry

Tobias Kaufmann*, Stefan Völker, Laura Gunesch and Andrea Kübler

Department for Psychology I, Institute for Psychology, University of Würzburg, Würzburg, Germany

Edited by:
Emanuel Donchin, University of South Florida, USA

Reviewed by:
Alireza Mousavi, Brunel University, UK
Michal Lavidor, Bar Ilan University, Israel

****Correspondence:***
Tobias Kaufmann, Department for Psychology I, University of Würzburg, Marcusstr. 9-11, 97070 Würzburg, Germany.
e-mail: tobias.kaufmann@uni-wuerzburg.de

Brain–computer interfaces (BCI) based on event-related potentials (ERP) allow for selection of characters from a visually presented character-matrix and thus provide a communication channel for users with neurodegenerative disease. Although they have been topic of research for more than 20 years and were multiply proven to be a reliable communication method, BCIs are almost exclusively used in experimental settings, handled by qualified experts. This study investigates if ERP–BCIs can be handled independently by laymen without expert support, which is inevitable for establishing BCIs in end-user's daily life situations. Furthermore we compared the classic character-by-character text entry against a predictive text entry (PTE) that directly incorporates predictive text into the character-matrix. $N = 19$ BCI novices handled a user-centered ERP–BCI application on their own without expert support. The software individually adjusted classifier weights and control parameters in the background, invisible to the user (auto-calibration). All participants were able to operate the software on their own and to twice correctly spell a sentence with the auto-calibrated classifier (once with PTE, once without). Our PTE increased spelling speed and, importantly, did not reduce accuracy. In sum, this study demonstrates feasibility of auto-calibrating ERP–BCI use, independently by laymen and the strong benefit of integrating predictive text directly into the character-matrix.

Keywords: brain–computer interface, user-centered, auto-calibration, predictive text entry, event-related potentials, assisitve technology, P300-Speller, ERP-BCI

INTRODUCTION

Event-related potentials (ERP) are brain signals in the human electroencephalogram (EEG) that occur in response to stimuli, e.g., visual stimuli (for review, see Polich, 2007). Farwell and Donchin (1988) suggested to utilize these potentials for brain–computer interface (BCI) based communication, particularly dedicated to patients with impaired motor control due to neurodegenerative disease (e.g., patients suffering from amyotrophic lateral sclerosis, ALS). Such ERP–BCI is often referred to as P300-BCI due to its usual dependence on the P300, a prominent positive ERP approximately 200–400 ms post-stimulus (Picton, 1992; Polich et al., 1997). It is elicited in an oddball-paradigm, in which participants focus their attention on a rare stimulus (the odd stimulus) among many frequent but irrelevant stimuli. In the visual ERP–BCI this oddball is typically realized by presenting letters and numbers of the alphabet on a computer screen and flashing these characters randomly. Users are then asked to focus their attention on flashings of the character they intend to spell while ignoring all other characters (e.g., counting the target flashes). After a certain number of character flashings, a classification algorithm detects ERPs in the recorded signal and thereof determines the intended character. Hence, healthy as well as severely impaired users can communicate words on a character-by-character basis (e.g., Nijboer et al., 2008).

Today, visual ERP–BCIs are most commonly used systems for BCI-based communication (for a review, see Kleih et al., 2011). It has been shown that they are reliably accurate (in healthy participants, e.g., Guger et al., 2009; Kleih et al., 2010; in patients, e.g., Sellers et al., 2006, 2010; Townsend et al., 2010) and robust for long-term use (Nijboer et al., 2008; Sellers et al., 2010). However, when considering daily use of a visual ERP–BCI system at current status of technology, some main barriers in terms of practicability and usability can be identified:

First, preparation effort for EEG acquisition with commonly used wet electrodes and herewith originating inconvenience (e.g., electrode gel, washing hair) were reported as being highly discomforting for patients using BCIs (Zickler et al., 2011). This issue is currently addressed by several manufacturers, e.g., by investigating dry electrodes. First prototypes have been tested for use in BCI (e.g., Grozea et al., 2011; Zander et al., 2011, for a review see Mak et al., 2011).

Second, since its first description in 1988, numerous modifications to the ERP–BCI have been well studied and led to many improvements e.g., in signal processing and classification (e.g., Krusienski et al., 2006, 2008; Blankertz et al., 2011, for reviews see Lotte et al., 2007; Kleih et al., 2011) and on the stimulus presentation side (e.g., Sellers et al., 2006; Guger et al., 2009; Salvaris and Sepulveda, 2009; Takano et al., 2009; Jin et al., 2010b; Townsend et al., 2010; Frye et al., 2011; Kaufmann et al., 2011; McFarland et al., 2011). Many of these modifications have been incorporated into today's BCI software systems and led to implementation of a plethora of parameters that allow for individual adjustment. However, being comprehensive and flexible for research purpose inevitably entails complexity and effort in user handling. Therefore, this constitutes a major problem when bringing BCI technology toward end-users (Zickler et al., 2009), as they and their care-givers may not be educated in handling of complex software as well as analysis and interpretation of acquired EEG data for calibrating the BCI system.

Third, communication speed and accuracy are low compared to existing solutions from the field of assistive technology (AT; Zickler et al., 2011). In parts this issue may be improved with further investigation of new signal processing methods and/or modifications of the paradigm. For example Kaufmann et al. (2011) recently investigated new stimuli, i.e., flashing characters with famous faces, which clearly outperformed classic character flashing in terms of communication speed and accuracy. Consequently, choice of appropriate stimuli may enhance bit rate. However, independent of all possible improvements, the commonly used ERP–BCI paradigm is necessarily limited to spelling on a character-by-character basis. Therefore, another attempt to foster bit rate is to integrate a predictive text entry (PTE) system into the BCI paradigm (Wolpaw et al., 2002). Besides the classic character-by-character text entry, Ryan et al. (2011) presented users with a separate window of numbered text suggestions, determined by a predictive text system. Users were provided with the possibility to select these numbered text predictions by focusing on their associated number in the matrix. For example, after spelling an "y," a first suggestion "your" would appear with number "1" among other suggestions in the separate window and by focusing on "1" in the character-matrix, users would in fact spell "your." We therefore further refer to this paradigm as "indirect paradigm." The authors reported that integration of such PTE system into the BCI effectively decreased the time needed to spell a sentence. A likewise indirect approach was proposed by Jin et al. (2010a) who adapted a Chinese cell-phone input system for BCI use to enable communication for languages that comprise too many characters for classic character-by-character text entry (e.g., Chinese, Japanese). In their system, seven word suggestions were presented that could be chosen indirectly by focusing on the numbers 1–7 in the matrix. Unfortunately, Ryan et al. (2011) reported that overall spelling accuracy was significantly decreased when using such indirect PTE paradigms, which was explained with higher workload and task demands (dual task interference). However, it remains unclear if this negative side effect is generalizable for all paradigms using predictive text system or if the higher workload is evoked by the indirect paradigm used in this study, i.e., selecting predictive text by focusing on its related number in the main matrix.

We addressed the latter two barriers by developing and evaluating a software solution, further referred to as optimized communication interface (OCI). It was specifically designed for use by laymen and thus reduces handling of the whole BCI application to a minimum of several button presses, thereby automatically controlling all other necessary steps in the background. Its core feature lies in an auto-calibration of the classifier that – with one button press – individually adjusts classifier weights to the user's brain signals. The system assesses quality of the signal recorded during a calibration session and decides if further calibration data is needed. Furthermore a PTE system was integrated as optional choice. Instead of numbering text suggestions and requiring the user to indirectly focus on the appropriate numbers (Ryan et al., 2011), we integrated predictive text directly into the matrix. This could decrease workload and task demand as users directly focus on the suggested word and its selection would be similar to selections of any other character in the matrix.

The herein presented study evaluated use of OCI by investigating (1) if auto-calibrating BCI systems can be used independently by laymen and (2) if predictive text incorporated into the speller matrix may better serve the goal of fast but highly accurate BCI-based communication than the classic character-by-character selection.

MATERIALS AND METHODS

OPTIMIZED COMMUNICATION INTERFACE (OCI) SOFTWARE

Implementation of OCI was based on the core modules of *BCI2000* (Schalk et al., 2004, Version 3, http://bci2000.org) and compiled with *Microsoft Visual Studio 2008* compiler. It can be grouped into three work stages: (1) modification of the graphical user interface, (2) implementation of an auto-calibration and (3) implementation of a stimulus presentation paradigm including a PTE system.

Graphical user interface

BCI2000's graphical user interface was modified using *Qt Designer* 4.7. Handling of the whole BCI application was reduced to one button and an optional check mark for switching on the PTE system (see **Figure 1**). When starting OCI, the button acts as a start button for the calibration session. With one button press, all parameters are automatically set in the background and the classic character-matrix appears. After a predefined calibration word was spelled, the matrix disappears and auto-calibration is performed in the background without user intervention. In case of successful calibration, the button automatically turns into a start button for the online, free-spelling mode. Otherwise it would act as a start button for recalibration of the system.

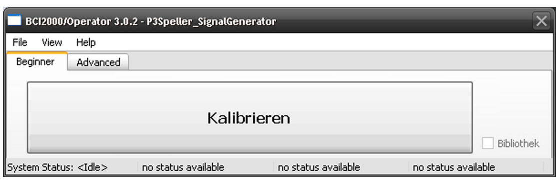


FIGURE 1 | Graphical user interface of OCI. The application is controlled with one button only. When starting OCI, the button displays "Kalibrieren" (German for "calibrate"). During the calibration phase it acts as a stop button, whereas after successful calibration it turns into a start button for the online mode. Optionally, a predictive text entry system is switched on by selecting the "Bibliothek" check mark (German for "library").

Auto-calibration

Auto-calibration was performed by automatically transferring the acquired data to *MATLAB* (The Mathworks, USA; version 2010b) using the *MATLAB Component Runtime* 4.14 to bridge between *C*++ and *MATLAB* code. A stepwise linear discriminant analysis (intervals of 800 ms post-stimulus, no spatial filtering) was performed based on data from all 12 EEG electrodes and thereof classifier weights were computed for the online sessions. Also, the number of sequences (NoS; a sequence consists of all rows and columns flashed once) used for flashing the characters in the online sessions was computed. Accuracy and stability of offline classification determined the outcome of the calibration session: (1) If accuracy was below 75% with 15 sequences, calibration session had to be performed again. (2) NoS used for online sessions was adjusted based on the stability of the classifier output, i.e., maximum accuracy remained stable after a given amount of data segments. In this case, NoS was defined as the minimum number of sequences to reach a stable offline accuracy maximum plus two sequences. E.g., if stable 100% accuracy were reached offline with seven sequences, NoS would have been nine during online sessions. (3) If no plateau was reached but classification accuracy was above 75%, NoS was set to 15. (4) The minimum NoS was restricted to eight to prevent from high error rates in long spelling sessions.

Note that auto-calibration was performed invisible to the user and all parameters for the online sessions were automatically adjusted without expert or user intervention.

Predictive text and stimulus presentation paradigm

First, we wrote an application that screened several prominent German internet pages (e.g., news pages) for words and counted their amount of occurrence. The word list was then sorted with decreasing occurrence and manually checked for non-sense words or words with internet specific content, e.g., name of the webpage holder. The final list comprised 82.616 words. This number includes various grammatical variations of the same word, e.g., plural forms.

Suggestion of predictive text and stimulus presentation matrix was implemented in Python 2.5 and connected to *BCI2000* via user datagram protocol. After each letter selection, the application screened the list downwards to update the six most likely words that were than presented in the first column of the speller matrix. If less than six words matched the previously entered letters, fewer up to none text suggestions were presented. At the beginning of new words, i.e., after selection of space, word deletion or after selection of a predictive word, the six most frequent words were presented (first six words of the list).

For the control condition, i.e., a matrix without PTE, an identical Python based stimulus presentation paradigm was used. Only it did not contain the PTE. **Figure 2** provides examples of both stimulus presentation matrices. All matrices were of size 6×6 and displayed on a computer screen approximately 70 cm distant from the participant (size: 22″; refresh rate: 60 Hz; resolution: 1680×1050; size of matrix: $40 \times 25\ \mathrm{cm}^2$).

As predictive text was directly integrated into the matrix it could be flashed in a similar manner than usually done in ERP-BCIs, i.e., in both conditions (PTE and character-by-character entry) the full 6×6 matrix was flashed row/column wise. Thus, selection of predictive text was also performed by counting each highlighting of the intended word to spell. The BCI system always chose the cell in the matrix with highest probability, which could be a selection of predictive text, single characters or commands (delete, space, escape) depending on the content of the chosen matrix cell (see **Figure 2**).

EXPERIMENTAL SCHEDULE

Electrodes were mounted by the experimenter. Thereafter, all participants handled the software fully on their own without expert support. Handling of OCI was explained in a condensed user manual (paper hardcopy) and with two short videos.

First, participants were required to start OCI and to start a calibration session where they had to spell the word "*BRAIN-POWER*." Each character was highlighted 30 times per letter to spell [NoS = 15; stimulus duration of 31.25 ms; inter-stimulus interval (ISI) of 125 ms]. This resulted in a final data set of 30×10 target and 150×10 non-target stimuli. From this data set, classifier weights were automatically computed in the background, invisible to the user.

Thereafter, participants twice spelled a German sentence comprising of 45 characters (nine words) using the automatically adjusted classifier weights and automatically determined NoS. Errors had to be corrected immediately such that the finalized sentence was 100% accurate. Then, spelling had to be terminated by selecting twice the "escape" function from the matrix, which brought them back to the main application window (necessary double selection constituted a safety mechanism of the "escape" function, to prevent unintended program termination due to selection error, i.e., selecting "escape" only once had no consequence). The inter-trial interval was 5 s to provide participants with enough time to perceive feedback on the spelled character/word and to screen all predictive text suggestions before the next stimulus sequence began. The same sentence was spelled once with PTE and once without, but both times with equal classifier weights and equal NoS to ensure comparability (within-subject design; calibration had to be performed only once). The order was randomized across participants.

After finalization of the spelling sessions, participants were asked to fill out questionnaires on handling of OCI and the PTE.

PARTICIPANTS

$N = 20$ healthy university students with normal or corrected to normal vision participated in the study. All participants were native German speakers and BCI novices. Due to equipment failure, data acquisition for $N = 1$ participant was aborted. The final sample comprised $N = 19$ participants (10 male, mean age 23.9 years, SD = 3.7, range 20–35). The experiment was conducted in accordance with standard ethical guidelines as defined by the Declaration of Helsinki (World Medical Association) and the European Council's Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine (Convention on Human Rights and Biomedicine). All participants gave signed informed consent prior to the experiment and received 8 € per hour for taking part in the study.

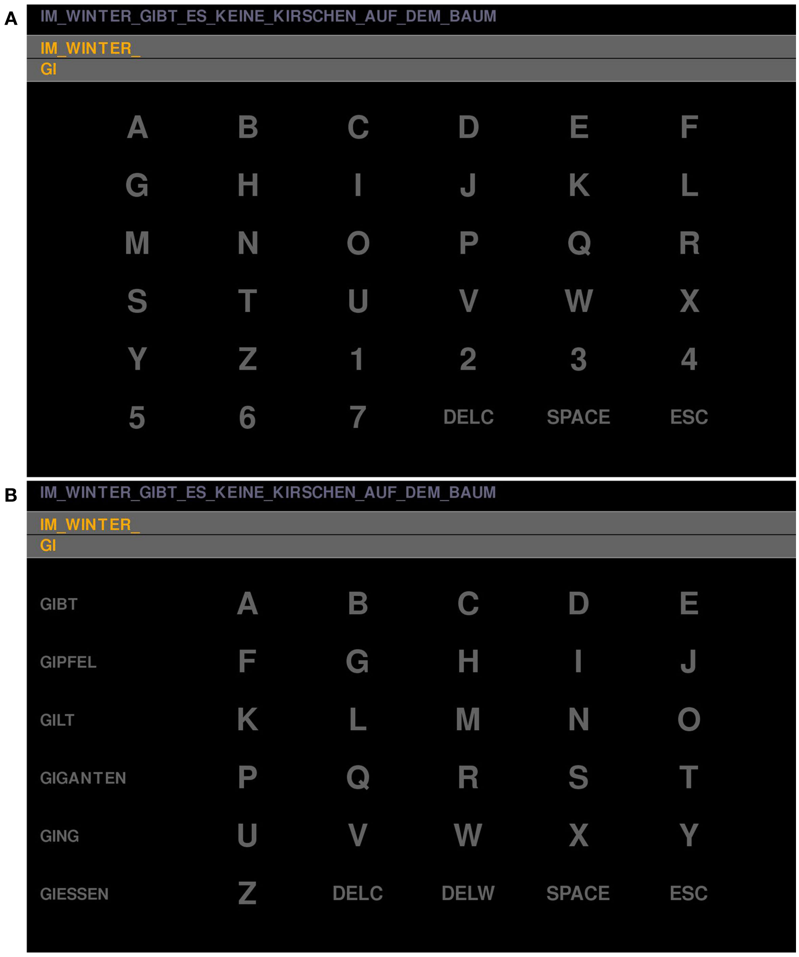


FIGURE 2 | Two paradigms were tested. (A) The classic character matrix consisting of single characters. Furthermore the matrix contains functions for deletion (delete character), space, and escape (finish writing). **(B)** Our new predictive text entry paradigm with predictive text included into the first column of the matrix to allow for direct selection of words. Also, the matrix contains functions for deletion (delete character, delete word), space, and escape.

DATA ACQUISITION

Electroencephalogram was acquired from 12 passive Ag/AgCl electrodes that were placed according to the guidelines of the American Electroencephalographic Society at positions Fz, FC1, FC2, C3, Cz, C4, P3, Pz, P4, O1, Oz, and O2. Furthermore an electrooculogram (EOG) was obtained from two horizontal (hEOG) and two vertical (vEOG) sintered Ag/AgCl electrodes. Both, EEG and EOG were recorded at a sampling rate of 256 Hz and amplified with a 16-channel g.USBamp amplifier (*g.tec Medical Engineering GmbH*, Austria).

ANALYSIS

The study aimed at investigating two independent aspects, i.e., feasibility of auto-calibrated BCI use and benefit of a new PTE paradigm.

Suitability of the auto-calibration was assessed by evaluating offline classifier accuracy as well as online spelling performance with the classic character-by-character-matrix condition. As auto-calibrated classifiers may build upon ocular artifacts due to missing control by the experimenter, we compared the auto-calibrated classifier accuracy to classifier accuracy after offline ocular artifact correction (Gratton et al., 1983). Benefit of our new PTE paradigm was assessed by comparing to the classic character-by-character-matrix the overall time needed to finish the sentence correctly, the amount of errors per selected item as well as achieved bit rate. Also, we evaluated results from questionnaires in terms of user satisfaction.

Besides computation of bit rate (as described e.g., in Wolpaw et al., 2000; Wolpaw et al., 2002; Serby et al., 2005) we computed a bit rate measure incorporating the information transferred with selections of whole words. PTE systems enhance bit rate in that with one selection many characters are communicated at once. Thus true bit rate is not only depending on the number of accurate selections made but also on the information transferred with each selection. Certainly such measure heavily depends on the length of the selected words. For example true bit rate would be five times higher when selecting a 10-character word instead of a two-character word, although bit rate in terms of selections/minute would be identical. Therefore true bit rate has to be interpreted with caution and is only provided to display the potential of PTE systems.

For offline analysis of ERPs, EEG data were filtered with 0.1 Hz high and 30 Hz low pass and corrected for ocular artifacts (Gratton et al., 1983). To avoid influence of overlapping ERPs due to the short ISI used in this study (ISI = 125 ms), only target stimuli with a minimum time distance of 469 ms (i.e., at least two non-target stimuli in between) were taken into account for ERP analysis. For statistical comparison, P300 peak amplitudes were obtained as maximum potentials between 200 and 300 ms at electrode Cz for both conditions.

Apart from ocular artifact correction with Brain Vision Analyser (Brain Products GmbH, Germany), all data analysis was performed using Matlab (The Mathworks, USA; version 2010b), and SPSS statistics (IBM, USA). Statistical comparison was performed using analysis of variance (ANOVA), except for not normally distributed data for which Wilcoxon-Test was used.

RESULTS

All participants were able to handle OCI on their own without expert support and to twice spell the predefined sentence correctly using the auto-calibrated classifier. In a forced-choice questionnaire all participants reported, that handling of the software was straightforward (average score: 3.84, with 1 = "not at all," 2 = "not really," 3 = "mostly," 4 = "completely"), that they would be able to manage the BCI application on their own in the future (3.84) and to explain it to others (3.74).

AUTO-CALIBRATION

For all participants, classifier weights were adjusted automatically after one calibration session and no repetitions were needed. In the consecutively scheduled online spelling sessions participants achieved overall good performance levels with regards to the length of the spelled sentence (average performance with classic character-by-character paradigm: 91.2%, SD = 7.8, range: 76–100% accuracy).

As auto-calibration may lead to classification of ocular artifacts, classifier accuracy of the auto-calibration as well as achieved stable NoS (see Auto-Calibration) were compared to accuracy and NoS after ocular artifact detection. No significant difference was found for the automatically selected NoS [$F(1, 18) = 0.156$, $p = .697$, $\eta^2_{partial} = 0.009$]. Also, offline classification accuracy did not differ before and after correction [With minimal NoS possible (NoS = 1): $F(1, 18) = 0.518$, $p = .481$, $\eta^2_{partial} = 0.028$; with minimal NoS used online (NoS = 8): $Z = -1.0$, $p = .317$]. Furthermore, we checked how many NoS were necessary to achieve a stable classifier accuracy of 100%. Difference was not significant before and after ocular artifact correction [$F(1, 18) = 0.010$, $p = .920$, $\eta^2_{partial} = 0.001$].

PREDICTIVE TEXT ENTRY

Figure 3 summarizes the overall time needed to finish each sentence correctly. In line with Ryan et al. (2011), PTE significantly decreased spelling time [$F(1, 18) = 95.74$, $p < .001$, $\eta^2_{partial} = 0.84$]. Average time needed to finish the task was reduced by factor 1.8 (12.4 min as compared to 22.3 min without PTE).

However, in contrast to the indirect paradigm by Ryan et al. (2011), accuracy did not decrease with our PTE paradigm [$F(1, 18) = 1.64$, $p = .216$, $\eta^2_{partial} = 0.084$]. Performance ranged from 100% accuracy to 76% (classic character-by-character condition) and to 74% respectively (PTE condition), see **Figure 4**. Bit rate (in terms of selections per minute) was comparably high for both conditions (character-by-character: $M = 15.1$ characters/min, SD = 5.6; PTE: $M = 15.7$ characters/min, SD = 5.7). **Figure 5** displays bit rate as well as true bit rate. True bit rate was higher for all participants, when using the PTE paradigm. For those participants that made many errors with this paradigm (P08, P17) benefit in terms of true bit rate was small. On the other hand, participants with high accuracy also show a high benefit. Average true bit rate for the sentence was $M = 12.0$ characters/min (SD = 2.7) with character-by-character text entry and $M = 20.6$ characters/min (SD = 5.3) with PTE. Please note that true bit rate heavily depends on the length of the selected words (see Analysis) and thus needs to be interpreted with caution. In this experiment, word length of the selected words was $M = 4.1$ characters on average (SD = 1.9; range: 2–8 characters).

Figure 6 displays the grand average ERPs for both conditions with most prominent P300 amplitude at electrode Cz around 246.5 ms (SD = 23.2; character-by-character text entry) and 250.0 ms (SD = 24.3; PTE). Grand average P300 amplitude was 3.55 μV (SD = 1.19) for character-by-character text entry and 3.44 μV (SD = 1.40) for PTE. No within-subject difference was found for P300 amplitudes between the conditions [$F(1, 18) = 0.72$, $p = .410$, $\eta^2_{partial} = 0.038$].

We further compared the amount of errors made when selecting a character against the amount of errors made when selecting a word or command. When using the classic character-by-character text entry, participants made $M = 2.2$ errors (SD = 2.2) on average for selecting a single character and $M = 4.2$ errors (SD = 4.8) for selecting a command. Error rates in the PTE condition were slightly lower (single character selection: $M = 1.3$, SD = 1.6; selection of command or word: $M = 1.6$, SD = 2.5). To write the sentence without any error, 10 commands (8× space, 2× escape) and 37 characters were necessary in the character-by-character text entry condition, whereas 11 commands and words and 16 characters were necessary in the PTE condition.

Evaluation of questionnaires revealed that all participants would prefer to use the PTE paradigm instead of character-by-character spelling, when exposed to BCI-based spelling in the future. Only one participant (participant 9) reported higher effort in selecting complete words as compared to single characters, which however was not reflected in his performance.

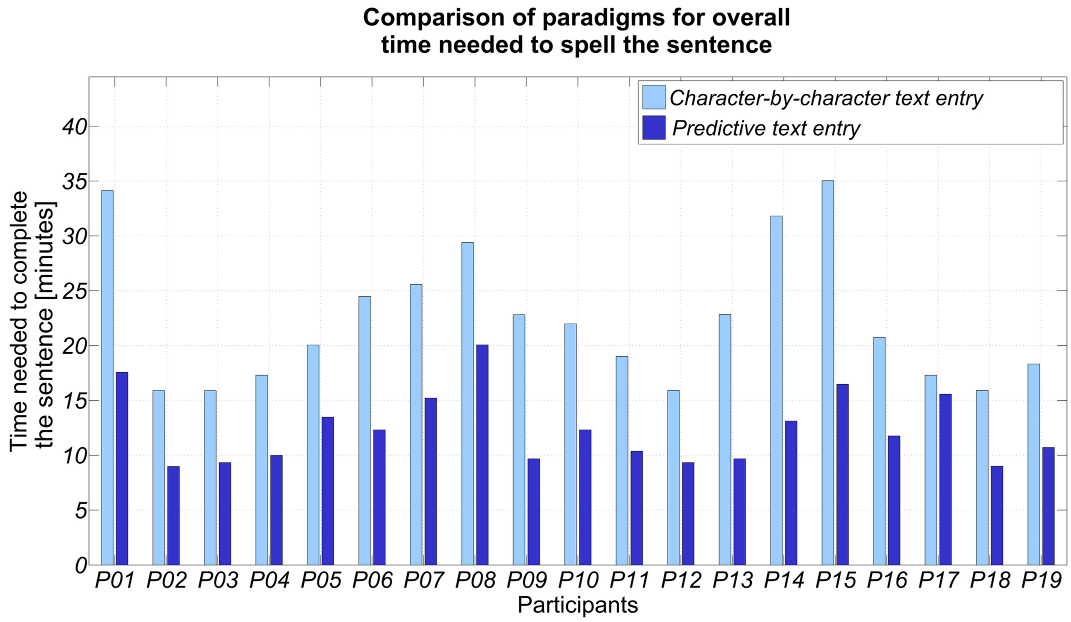


FIGURE 3 | Overall time needed to copy spell the sentence (including error correction and escape of spelling mode after finalization of 100% accurate sentence): within-subject comparison between classic character-by-character text entry and predictive text entry paradigm.

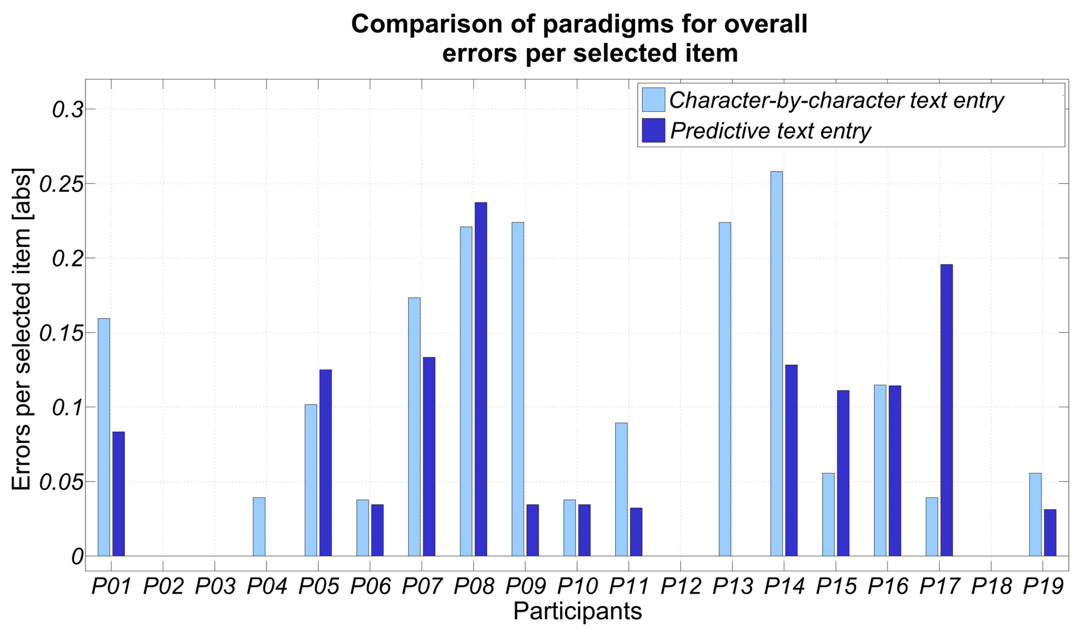


FIGURE 4 | Average number of errors per selected item: within-subject comparison between classic character-by-character text entry and predictive text entry paradigm. As the number of items selected to finish the sentence varied between participants and conditions (e.g., due to error correction or selection of predictive text) no absolute amount of errors is presented but instead the amount of errors per selected item.

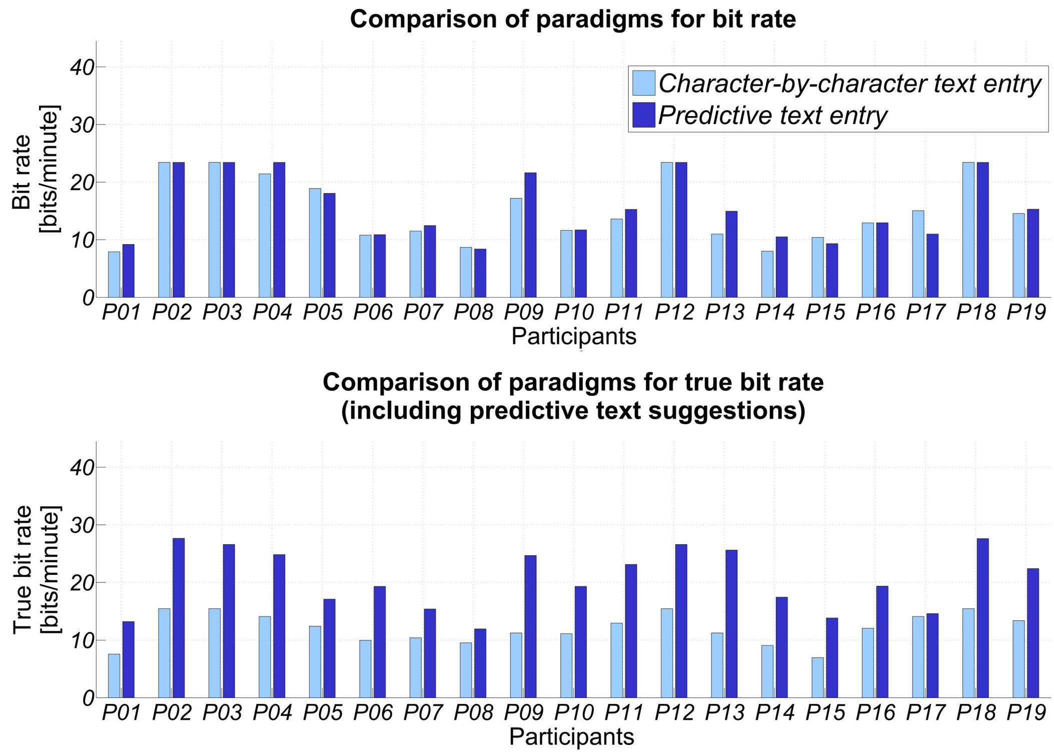


FIGURE 5 | Bit rate (in terms of selections per time unit) and true bit rate (in terms of communicated characters per time unit): within-subject comparison between classic character-by-character text entry and predictive text entry paradigm.

DISCUSSION

This study provides evidence, that current BCI technology can be automatized and operated by BCI novices through a user-centered, easy-to-use graphical user interface. This is inevitable when bringing BCIs out of the laboratory environment to end-user's homes. Zickler et al. (2009) assessed user needs and requirements from $N = 77$ AT users and reported "functionality," "possibility of independent use," and "easiness of use" as most important requirements. The latter was one of the main reasons for dissatisfaction with current AT solutions. Importantly, our results from handling of OCI indicate high satisfaction with both, "easiness of use" and "possibility of independent use." Furthermore, "functionality" was high as none of the auto-calibration sessions had to be performed twice and all participants were able to correctly spell the sentences. Overall spelling accuracy was high and none of the participants performed with lower than 70% accuracy – an accuracy level previously described as minimum level for communication (Kübler et al., 2001).

When considering adoption of a new AT solution, "functionality/effectiveness" was identified as the major critical factor for potential end-users (Zickler et al., 2009). This study suggested integration of predictive text directly into the matrix and proofed high benefit in terms of spelling speed (effectiveness) without loss of accuracy (functionality). As can be particularly seen from the true bit rate of participants with lower accuracy in the PTE paradigm, a high performance level is inevitable to clearly benefit from a PTE system. The unaffected P300 amplitudes in our study show that when integrating predictive text directly into the matrix workload may remain low, as predictive text is selected in a similar manner than single characters. The drop in performance reported by Ryan et al. (2011) may thus indeed be attributed to higher workload and dual task interference in the indirect PTE paradigm. The remaining performance level in our study, as well as the heavily decreased time needed to spell a sentence may well explain why none of the participants reported to prefer the character-by-character over the PTE paradigm. Advantageously, daily life application with such PTE system could easily be equipped with algorithms for grammar recognition and learning of new words.

LIMITATIONS AND FUTURE EXPERIMENTATION

To reduce the amount of errors during spelling of long sentences, we restricted the automatically adjusted NoS to a minimum of eight sequences. This limit was set on basis of previous experience. Apparently, there is a trade-off, as low NoS increases the probability of errors due to a low number of trials to average, whereas high NoS may entail errors due to increased duration of spelling time. We speculate that for some participants NoS could be further reduced. This could also be realized with recently introduced

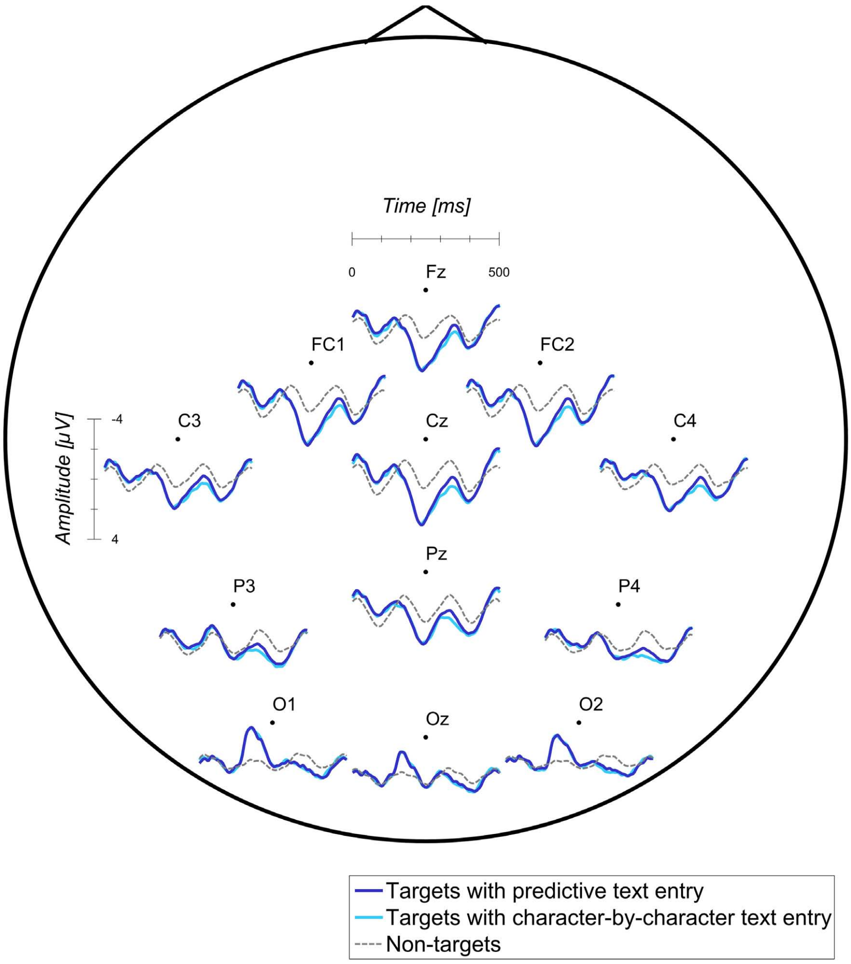


FIGURE 6 | Grand average event-related potentials: within-subject comparison between classic character-by-character text entry and predictive text entry paradigm.

dynamic stopping methods. (e.g., Lenhardt et al., 2008; Höhne et al., 2010; Liu et al., 2010; Schreuder et al., 2011b; for comparison of methods see Schreuder et al., 2011a). Instead of flashing characters with a fixed NoS, dynamic stopping algorithms adjust the NoS in every trial, i.e., stimulus presentation stops as soon as probability of correct target classification reaches a certain level.

This study aimed at investigating if spelling at high accuracy level is possible using an auto-calibrated classifier. Comparison to manually calibrated classifier performance needs separate investigation. Manual calibration may still be valuable for some cases (e.g., for exclusion of bad channels or detection of equipment failure). Auto-calibration may also be further equipped with mechanisms for detecting such cases. Our results are promising in that high accuracy could be achieved in all participants using auto-calibration and none of the participants needed recalibration of the system.

We assume that our validation results from healthy users may well transfer to end-users as we cautiously addressed their major concerns (Zickler et al., 2009). However, Zickler et al. (2010) found differences between end-users and care-givers with regards to their expectations of new AT solutions – differences which may certainly also exist in our healthy sample. Thus, it is inevitable to develop and validate new BCI technology in close discussion with them. Future experimentation therefore needs to validate our results in an end-user sample.

CONCLUSION

In sum, this study specifically addressed critical aspects of ERP–BCI-based communication reported by end-users, i.e., "functionality/effectiveness," "easiness of use," and "possibility of independent use." Our results proofed feasibility of expert independent BCI-based communication using auto-calibration and underlined the strong benefit of predictive text directly integrated into the spelling matrix.

ACKNOWLEDGMENTS

We thank Ruben Real for his support. This work is supported by the European ICT Programme Project FP7-224631. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein. Open access publication was funded by the German Research Foundation (DFG) and the University of Würzburg in the funding programme Open Access Publishing."

REFERENCES

Blankertz, B., Lemm, S., Treder, M., Haufe, S., and Müller, K.-R. (2011). Single-trial analysis and classification of ERP components – a tutorial. *Neuroimage* 56, 814–825.

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

Frye, G. E., Hauser, C. K., Townsend, G., and Sellers, E. W. (2011). Suppressing flashes of items surrounding targets during calibration of a P300-based brain-computer interface improves performance. *J. Neural Eng.* 8, 025024.

Gratton, G., Coles, M. G., and Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalogr. Clin. Neurophysiol.* 55, 468–484.

Grozea, C., Voinescu, C. D., and Fazli, S. (2011). Bristle-sensors – low-cost flexible passive dry EEG electrodes for neurofeedback and BCI applications. *J. Neural Eng.* 8, 025008.

Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., Gramatica, F., and Edlinger, G. (2009). How many people are able to control a P300-based brain-computer interface (BCI)? *Neurosci. Lett.* 462, 94–98.

Höhne, J., Schreuder, M., Blankertz, B., and Tangermann, M. (2010). Two-dimensional auditory p300 speller with predictive text system. *Conf. Proc. IEEE Eng. Med. Biol. Soc. 2010*, 4185–4188.

Jin, J., Allison, B. Z., Brunner, C., Wang, B., Wang, X., Zhang, J., Neuper, C., and Pfurtscheller, G. (2010a). P300 Chinese input system based on Bayesian LDA. *Biomed. Tech. (Berl.)* 55, 5–18. doi:10.1515/BMT.2010.003

Jin, J., Horki, P., Brunner, C., Wang, X., Neuper, C., and Pfurtscheller, G. (2010b). A new P300 stimulus presentation pattern for EEG-based spelling systems. *Biomed. Tech. (Berl.)* 55, 203–210. doi:10.1515/BMT.2010.029

Kaufmann, T., Schulz, S. M., Grünzinger, C., and Kübler, A. (2011). Flashing characters with famous faces improves ERP-based brain-computer interface performance. *J. Neural Eng.* 8, 056016.

Kleih, S. C., Kaufmann, T., Zickler, C., Halder, S., Leotta, F., Cincotti, F., Aloise, F., Riccio, A., Herbert, C., Mattia, D., and Kübler, A. (2011). Out of the frying pan into the fire – the P300-based BCI faces real-world challenges. *Prog. Brain Res.* 194, 27–46.

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

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

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2006). A comparison of classification techniques for the P300 Speller. *J. Neural Eng.* 3, 299–305.

Kübler, A., Neumann, N., Kaiser, J., Kotchoubey, B., Hinterberger, T., and Birbaumer, N. P. (2001). Brain-computer communication: self-regulation of slow cortical potentials for verbal communication. *Arch. Phys. Med. Rehabil.* 82, 1533–1539.

Lenhardt, A., Kaper, M., and Ritter, H. J. (2008). An adaptive P300-based online brain-computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* 16, 121–130.

Liu, T., Goldberg, L., Gao, S., and Hong, B. (2010). An online brain-computer interface using non-flashing visual evoked potentials. *J. Neural Eng.* 7, 036003.

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., and Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain-computer interfaces. *J. Neural Eng.* 4, R1–R13.

Mak, J. N., Arbel, Y., Minett, J. W., McCane, L. M., Yuksel, B., Ryan, D., Thompson, D., Bianchi, L., and Erdogmus, D. (2011). Optimizing the P300-based brain-computer interface: current status, limitations and future directions. *J. Neural Eng.* 8, 025003.

McFarland, D. J., Sarnacki, W. A., Townsend, G., Vaughan, T., and Wolpaw, J. R. (2011). The P300-based brain-computer interface (BCI): effects of stimulus rate. *Clin. Neurophysiol.* 122, 731–737.

Nijboer, F., Sellers, E. W., Mellinger, J., Jordan, M. A., Matuz, T., Furdea, A., Halder, S., Mochty, U., Krusienski, D. J., Vaughan, T. M., Wolpaw, J. R., Birbaumer, N., and Kübler, A. (2008). A P300-based brain – computer interface for people with amyotrophic lateral sclerosis. *Clin. Neurophysiol.* 119, 1909-1916.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* 9, 456–479.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148.

Polich, J., Alexander, J. E., Bauer, L. O., Kuperman, S., Morzorati, S., O'Connor, S. J., Porjesz, B., Rohrbaugh, J., and Begleiter, H. (1997). P300 topography of amplitude/latency correlations. *Brain Topogr.* 9, 275–282.

Ryan, D. B., Frye, G. E., Townsend, G., Berry, D. R., Mesa-G, S., Gates, N. A., and Sellers, E. W. (2011). Predictive spelling with a P300-based brain-computer interface: increasing the rate of communication. *Int. J. Hum. Comput. Interact.* 27, 69–84.

Salvaris, M., and Sepulveda, F. (2009). Visual modifications on the P300 speller BCI paradigm. *J. Neural Eng.* 6, 046011.

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., and Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans. Biomed. Eng.* 51, 1034–1043.

Schreuder, M., Höhne, J., Treder, M., Blankertz, B., and Tangermann, M. (2011a). Performance optimization of ERP-based BCIs using dynamic stopping. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2011, 4580–4583.

Schreuder, M., Rost, T., and Tangermann, M. (2011b). Listen, you are writing! speeding up online spelling with a dynamic auditory BCI. *Front. Neurosci.* 5:112. doi:10.3389/fnins.2011.00112

Sellers, E. W., Kübler, A., and Donchin, E. (2006). Brain-computer interface research at the University of South Florida Cognitive Psychophysiology Laboratory: the P300 Speller. *IEEE Trans. Neural Syst. Rehabil. Eng.* 14, 221–224.

Sellers, E. W., Vaughan, T. M., and Wolpaw, J. R. (2010). A brain-computer interface for long-term independent home use. *Amyotroph. Lateral Scler.* 11, 449–455.

Serby, H., Yom-Tov, E., and Inbar, G. F. (2005). An improved P300-based brain-computer interface. *IEEE Trans. Rehabil. Eng.* 13, 89–98.

Takano, K., Komatsu, T., Hata, N., Nakajima, Y., and Kansaku, K. (2009). Visual stimuli for the P300 brain–computer interface: a comparison of white/gray and green/blue flicker matrices. *Clin. Neurophysiol.* 120, 1562–1566.

Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C. K., Schwartz, N. E., Vaughan, T. M., Wolpaw, J. R., and Sellers, E. W. (2010). A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clin. Neurophysiol.* 121, 1109–1120.

Wolpaw, J. R, Birbaumer, N., Heetderks, W. J., McFarland, D. J., Peckham, P. H., Schalk, G., Donchin, E., Quatrano, L. A., Robinson, C. J., and Vaughan, T. M. (2000). Brain-computer interface technology: a review of the first international meeting. *IEEE Trans. Rehabil. Eng.* 8, 164–173.

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

Zander, T. O., Lehne, M., Ihme, K., Jatzev, S., Correia, J., Kothe, C., Picht, B., and Nijboer, F. (2011). A dry EEG-system for scientific research and brain-computer interfaces. *Front. Neurosci.* 5:53. doi:10.3389/fnins.2011.00053

Zickler, C., Di Donna, V., Kaiser, V., Al-Khodairy, A., Kleih, S. C., and Kübler, A. (2009). "BCI applications for people with disabilities: defining user needs and user requirements," in *Assistive Technology from Adapted Equipment to Inclusive Environments, AAATE. 25 Assistive Technology Research Series.* eds P. L. Emiliani, L. Burzagli, A. Como, F. Gabbanini, and A.-L. Salminen (Amsterdam: IOS Press), 185–189.

Zickler, C., Kaiser, V., Al-Khodairy, A., Kleih, S. C., Kübler, A., Mattia, D., Mongardi, S., Rohm, M., Rupp, R., Staiger-Sälzer, P., and Hoogerwerf, E. J. (2010). "BCI-applications: requirements of disabled end-users and professional users," in *Poster presented at the 1st TOBI Workshop 2010*, Graz.

Zickler, C., Riccio, A., Leotta, F., Hillian-Tress, S., Halder, S., Holz, E., Staiger-Sälzer, P., Hoogerwerf, E. J., Desideri, L., Mattia, D., and Kübler, A. (2011). A brain-computer interface as input channel for a standard assistive technology software. *Clin. EEG Neurosci.* 42, 236–244.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 February 2012; accepted: 30 April 2012; published online: 23 May 2012.

Citation: Kaufmann T, Völker S, Gunesch L and Kübler A (2012) Spelling is just a click away – a user-centered brain–computer interface including auto-calibration and predictive text entry. Front. Neurosci. ***6****:72. doi: 10.3389/fnins.2012.00072*

This article was submitted to Frontiers in Neuroprosthetics, a specialty of Frontiers in Neuroscience.

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

Abstract: Currently, the event-related potential (ERP)-based spelling device, often referred to as P300-Speller, is the most commonly used brain-computer interface (BCI) for enhancing communication of patients with impaired speech or motor function. Among numerous improvements, a most central feature has received little attention, namely optimizing the stimulus used for eliciting ERPs. Therefore we compared P300-Speller performance with the standard stimulus (flashing characters) against performance with stimuli known for eliciting particularly strong ERPs due to their psychological salience, i.e. flashing familiar faces transparently superimposed on characters. Our results not only indicate remarkably increased ERPs in response to familiar faces but also improved P300-Speller performance due to a significant reduction of stimulus sequences needed for correct character classification. These findings demonstrate a promising new approach for improving the speed and thus fluency of BCI-enhanced communication with the widely used P300-Speller.

Link to publication source: Journal of Neural Engineering, IOP Publishing

Date of written permission: March 11, 2013.

IOP Publishing Journal of Neural Engineering
J. Neural Eng. **8** (2011) 056016 (10pp) doi:10.1088/1741-2560/8/5/056016

Flashing characters with famous faces improves ERP-based brain–computer interface performance

T Kaufmann, S M Schulz[1], C Grünzinger and A Kübler

Department of Psychology I, University of Würzburg, Marcusstr. 9-11, 97070 Würzburg, Germany

E-mail: schulz@psychologie.uni-wuerzburg.de

Received 20 April 2011
Accepted for publication 4 August 2011
Published 20 September 2011
Online at stacks.iop.org/JNE/8/056016

Abstract
Currently, the event-related potential (ERP)-based spelling device, often referred to as P300-Speller, is the most commonly used brain–computer interface (BCI) for enhancing communication of patients with impaired speech or motor function. Among numerous improvements, a most central feature has received little attention, namely optimizing the stimulus used for eliciting ERPs. Therefore we compared P300-Speller performance with the standard stimulus (flashing characters) against performance with stimuli known for eliciting particularly strong ERPs due to their psychological salience, i.e. flashing familiar faces transparently superimposed on characters. Our results not only indicate remarkably increased ERPs in response to familiar faces but also improved P300-Speller performance due to a significant reduction of stimulus sequences needed for correct character classification. These findings demonstrate a promising new approach for improving the speed and thus fluency of BCI-enhanced communication with the widely used P300-Speller.

1. Introduction

Brain–computer interfaces (BCIs) provide an alternative communication channel that does not depend on muscular control and thus, have been found to be particularly useful for patients diagnosed with amyotrophic lateral sclerosis (ALS) who experience a gradual loss of control over voluntary muscular movement due to motor neuron degeneration (Birbaumer *et al* 1999, Kübler *et al* 2001a, Nijboer *et al* 2008).

The ERP-based spelling device, often referred to as P300-Speller, is one of the most commonly used BCIs with healthy participants as well as patients (Farwell and Donchin (1988), Nijboer *et al* (2008); for a review see Kleih *et al* (2011)). About 250 to 400 ms after stimulus onset a marked positive potential deflection can be observed in the event-related potential (ERP, i.e. the P300, Sutton *et al* (1965), Polich *et al* (1997); for a review see Picton (1992) and Polich (2007)). In BCI paradigms its latency can be even as early as 200 ms (Sellers *et al* 2010, Kleih *et al* 2010). P300 peak amplitudes can be recorded best at parietal and central electrodes (Picton 1992).

[1] Author to whom any correspondence should be addressed.

Strong P300s are reliably elicited in the oddball paradigm, in which participants are required to attend to rare (odd) target stimuli in a series of many irrelevant stimuli. In the most common version of the P300-Speller, the participant attends a target character in a matrix consisting of the letters of the alphabet and digits. Then the rows and columns of the matrix are flashed consecutively in random order (row column paradigm (RCP), see also Farwell and Donchin (1988)). An increased P300 will follow flashing of the particular column and row where the attended character is located.

Although healthy participants produce highly accurate results with the P300-Speller (Guger *et al* 2009, Kleih *et al* 2010), accuracy is still an issue in patients, because P300 amplitudes are typically lower than in healthy controls (Hanagasi *et al* 2002). For efficient communication maximizing the speed of character selection without increasing error rates is the primary goal. Many researchers have addressed this issue by improving (1) the quality of signal processing and the development of novel classification techniques (e.g. Krusienski *et al* (2006), (2008), Blankertz *et al* (2011); for a review see Lotte *et al* (2007)) and (2) procedural aspects of the oddball paradigm. Examples for the second approach are studies comparing how stimulus rate (McFarland *et al* 2011), inter-stimulus interval (ISI) and matrix size (Sellers *et al* 2006a), character size and inter-character distance (Salvaris and Sepulveda 2009), background color (Salvaris and Sepulveda 2009), chromatic differences in the flashing pattern (Takano *et al* 2009) and the flashing pattern itself (e.g. Townsend *et al* (2010), Frye *et al* 2011, Guger *et al* (2009), Jin *et al* (2010)) affect the accuracy level.

Altogether these studies suggest several modifications to the paradigm that may help to increase classification rates, bit rates or user acceptance, or both of the P300-Speller. Optimizing the core feature of the P300-Speller, that is reliably eliciting an ERP by means of flashing items in a matrix, has almost never been scrutinized. To our knowledge, only one study has tried to improve spelling performance by changing the immediate properties of this stimulus (Martens *et al* 2009). Instead of flashing characters, gray rectangles were flipped horizontally to elicit ERPs. This so called FLIP stimulus was less susceptible to refractory effects in ERPs. Thus, the bit rate may be slightly increased by decreasing the target-to-target intervals.

A more direct approach for improving classification accuracy is to optimize the immediate properties of the stimulus used for eliciting ERPs to increase the signal-to-noise ratio in the electro-encephalogram (EEG). This would allow for improved classification based on fewer ERPs, that is fewer stimulus repetitions (sequences). The latter would increase the bit rate and thus, enhance the fluency of communication. This goal may be accomplished by introducing a new stimulus type for flashing the characters which is known to elicit particularly strong ERPs. Ideal candidates for such a stimulus may be familiar faces. Numerous studies in the field of human face processing have revealed that the visual perception of familiar faces strongly involves several ERPs, specifically the N170, P300 and N400f. The N170 occurs between 130 and 200 ms post-stimulus and appears to be a face-specific ERP (Bentin *et al* 1996, Eimer 2000). It is assumed that the N170 is involved in rapid perception of faces but not in face recognition (Eimer 2000). Familiar faces further elicit a face-specific N400, the N400f, between 300 and 500 ms post-stimulus, which is located on parietal and central electrode sites (Eimer 2000). The recognition of familiar faces further involves a late positive component (P300/LPC, Henson *et al* (2003)). Classification in an ERP-BCI is not necessarily limited to the P300 component and other ERPs may also contribute (e.g. Bianchi *et al* (2010), Treder and Blankertz (2010), Allison and Pineda (2003, 2006)). Thus, presentation of familiar faces, which strongly involves different ERP components, may be exploited for improving classification. In particular using faces well known to everybody in a given culture should lead to strong and relatively stable effects across individuals.

Thus, we investigated whether familiar faces of famous individuals may serve as applicable stimuli for flashing items in the character matrix of the classic P300-Speller. We hypothesized that eliciting ERPs involved in face processing would lead to better classification rates as compared to simple letter highlighting, due to better signal determination within the response window.

2. Methods

2.1. Participants

Participants were $N = 21$ university students with no known neurological disorders and normal or corrected-to-normal vision. Due to noisy EEG, data from 1 participant were discarded. The remaining sample comprised $N = 20$ participants (10 men; mean age 24.6 years, $SD = 4.55$, range 19–41). All were BCI novices and signed informed consent prior to the study. The experiment was conducted in accordance with standard ethical guidelines as defined by the Declaration of Helsinki (World Medical Association) and the European Council's Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine (Convention on Human Rights and Biomedicine).

2.2. Equipment and data acquisition

EEG was recorded from 12 passive Ag/AgCl electrodes, as shown in figure 1, with mastoid ground and reference. Electro-oculogram (EOG) was acquired from 2 vertical (vEOG) and 2 horizontal (hEOG) passive Ag/AgCl sintered electrodes. All signals were amplified using a 16-channel g.USBamp amplifier (*g.tec Medical Engineering GmbH*, *Austria*) and recorded at a sampling rate of 512 Hz. Data were collected with the software tool BCI2000 (Schalk *et al* 2004).

2.3. The spelling paradigm

The spelling paradigm was implemented as a multi-thread application in Python 2.5 and connected to BCI2000 via user datagram protocol. Spelling characters were presented in a 6×6 matrix and rows and columns of the matrix were flashed consecutively in random order (RCP). We compared the classic character flashing (CF) to a novel face flashing (FF) condition. To ensure equal timing of stimulus presentation, both conditions were implemented and controlled with the same hardware and software implementation. The CF condition defaulted to the traditional approach (i.e. as originally described by Farwell and Donchin (1988)). In the FF condition, we overlaid rows and columns of the characters and digit matrix with well-known pictures of famous faces (Albert Einstein or Ernesto 'Che' Guevara). Pictures were semi-transparent to allow for uninterrupted focusing on the target letter while flashing the face stimuli. In addition to the goal of increasing the psychological salience of the stimulus, facial stimuli may differ from flashing characters also with regard to physical parameters such as brightness, contrast or chromatic differences. Chromatic differences, for example, were recently found to influence performance in the P300-BCI (Takano *et al* 2009). Therefore, we also implemented a control condition using a random reorganization of all pixels of the FF pictures, resulting in meaningless images with the exact same brightness, contrast and chromatic properties as compared to the FF condition. Consequently, the remaining difference between this pixelated face flashing (PFF) condition and FF is the configuration and semantic properties of the picture that

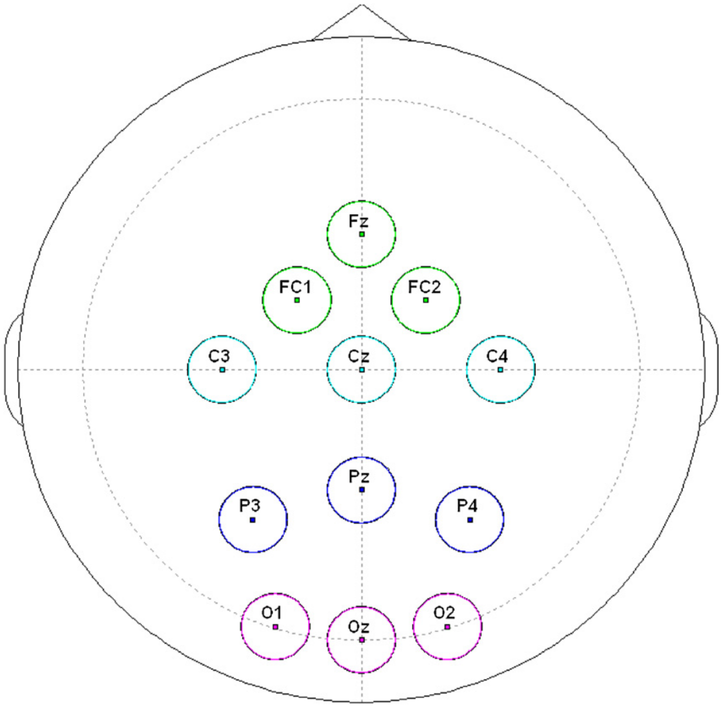


Figure 1. Electrode setup consisting of 12 passive Ag/AgCl electrodes. Locations were Fz, FC1, FC2, C3, Cz, C4, P3, Pz, P4, O1, Oz and O2.

is a prominent face. In total this resulted in five conditions: one CF condition, two FF conditions (Einstein versus 'Che' Guevara) and two PFF conditions (pixelated Einstein versus pixelated 'Che' Guevara). Figure 2 shows examples of the CF, FF and PFF conditions.

2.4. Procedure

Participants were seated comfortably in a chair approximately 70 cm distant from a computer screen (size: 19″, resolution: 1280 × 1024 pixel, refresh rate: 60 Hz, size of presented matrix: 30 × 20 cm^2). After informed consent was signed, they were instructed to focus on the target character, avoid blinking during stimulus presentation and to count the number of stimulus flashes. A trial necessary to select a target (i.e. letter to spell) comprised 15 flashes of each row and each column (sequences). The speller software was set to default values for the ISI of 125 ms and a flash duration of 31.25 ms, resulting in 33.33 ms real flash duration and 133.33 ms ISI at 60 Hz screen refresh rate. All participants spelled the word 'BRAIN' in each of the five conditions. The order of condition was permuted across participants to control for potential habituation effects. This procedure was repeated once, resulting in 2 × 5 runs in total. No feedback on the correctness of spelling was provided to the participants during any of the runs.

To avoid novelty effects, prior to each run participants were familiarized for 15 s with the stimulus used for flashing characters, i.e. the face picture in an FF condition. After the run, they rated on a 10-point Likert-like scale from 0 (not familiar at all) to 10 (extremely familiar) how familiar the stimulus appeared to them. Scales were presented with E-Prime® (*Psychology Software Tools*, *USA*).

2.5. Artifact rejection and data processing

EEG data were filtered with 0.1 Hz high and 30 Hz low pass. EEG was corrected for ocular artifacts with both hEOG and vEOG recordings using an algorithm introduced by Gratton *et al* (1983). These preprocessing steps were performed in Brain Vision Analyzer® (*Brain Products GmbH*, *Germany*). Data were then exported to Matlab® (*The MathWorks*, *USA*) for all further data analysis including signal classification. ERP amplitudes were determined as the maximum (positive potentials) or minimum (negative potentials) voltage in a defined time window.

This study used short ISIs of 125 ms, which bears the risk of overlapping ERPs. Regarding BCI performance, however, it could be demonstrated that short (175 ms) ISIs yield better performance results as compared to longer (350 ms) ISIs (Sellers *et al* 2006a). Furthermore, immediate consecutive flashing of the same target is rare due to the properties of the oddball paradigm. The reason for and advantage of short ISIs is rapid character selection which fosters the applicability of the P300-Speller. However, as we were not only interested in the effect of stimulus presentation on classification accuracy, but also on ERP amplitudes and latencies in different conditions, only target stimuli separated

J. Neural Eng. **8** (2011) 056016 T Kaufmann *et al*

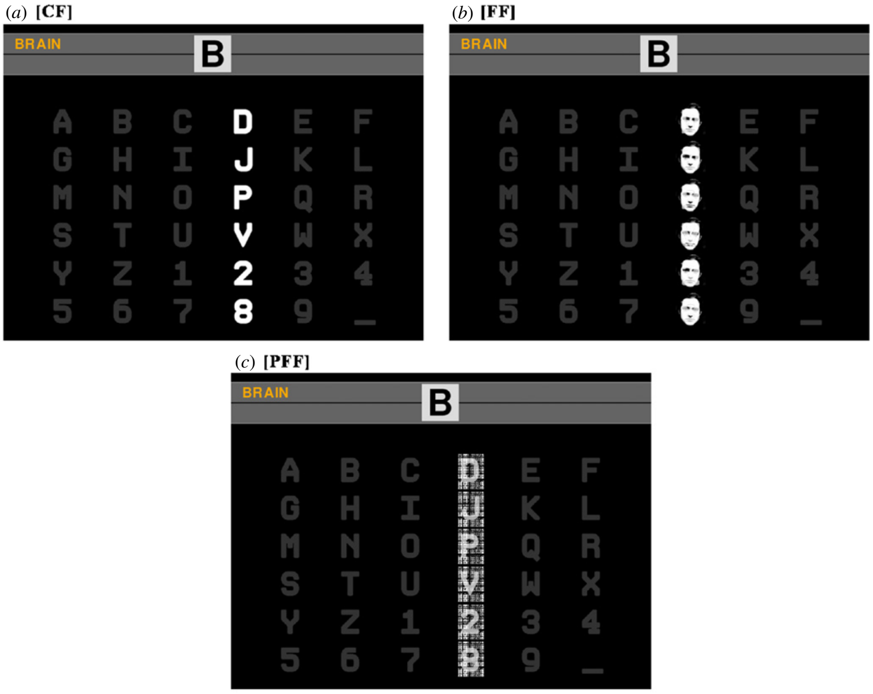


Figure 2. The different conditions of the paradigm. (*a*) The classic P300-Speller (CF), (*b*) the characters with flashing familiar faces (FF) and (*c*) the control condition with all pixels of the face picture after spatial randomization (PFF). Both FF and PFF stimuli were presented semi-transparent such that the characters were still visible. In the experiment we used famous faces which are not shown in this figure due to print license.

by more than two non-target stimuli were included for offline analysis of ERPs, resulting in a minimum time distance of 466 ms between two target stimuli. This correction was performed for ERP analysis only. The classification results comprise the full data set without this correction, thus reflecting the real P300-Speller condition. Classification of target versus non-target stimuli was performed using stepwise linear discriminant analysis (Krusienski *et al* 2006).

2.6. Statistical analysis

Differences in the number of sequences needed for offline classification between the CF and FF conditions and differences in ERP amplitudes and determination coefficients between CF, FF and PFF conditions were statistically validated using repeated measures analysis of variance (ANOVA) (with 'condition' as within subject factor). *Post-hoc* comparisons were performed with Tukey–Kramer tests.

3. Results

All participants completed all scheduled runs. The familiarity of the face stimuli as rated on a 10-point Likert-like scale was high ($M = 8.5$, $SD = 0.7$ for Einstein (FF); $M = 8.2$, $SD = 0.9$ for Che Guevara (FF)). In contrast, the familiarity of the pixelated control was rated low ($M = 2.4$, $SD = 2.3$ for pixelated Einstein (PFF); $M = 2.6$, $SD = 2.0$ for pixelated Che Guevara (PFF)).

3.1. Classification results

In the offline analysis a classifier was trained for the CF and FF conditions. As expected classifier accuracy was increased when faces were used as flashing stimuli (FF) as compared to the classic approach (CF), see figure 3. In the FF conditions, on average fewer sequences ($M = 2.9$, $SD = 1.8$ (Einstein) and $M = 3.0$, $SD = 1.7$ (Che Guevara)) were necessary for achieving the goal of stable 100% classifier accuracy as compared to the classic flashing paradigm ($M = 5.3$, $SD = 3.0$). A repeated measures ANOVA (within factor condition: CF, FF Einstein, FF Che Guevara) confirmed high significance ($F(2, 38) = 7.69$, $p < 0.002$, $\eta^2_{\text{partial}} = 0.29$). Besides the 100% accuracy level (maximum accuracy) we assessed two further levels of accuracy, as comprehensibility of communication is not only possible with the maximum accuracy level. A level of $\geqslant$70% may be regarded as a minimum level of communication (Kübler *et al* 2001b, Kübler and Birbaumer 2008), although communication accuracy not only depends on the amount of falsely spelled characters but also on the influence of the error on the word meaning. We therefore use the $\geqslant$70% as a reference for comparison to demonstrate the potential for improvement of communication with our novel approach. This $\geqslant$70% level was achieved with fewer stimulus sequences in the FF condition ($F(2, 38) = 4.95$, $p < 0.012$, $\eta^2_{\text{partial}} = 0.20$), for the majority of participants already after the first sequence. Secondly, the same but even more pronounced effect was found for a

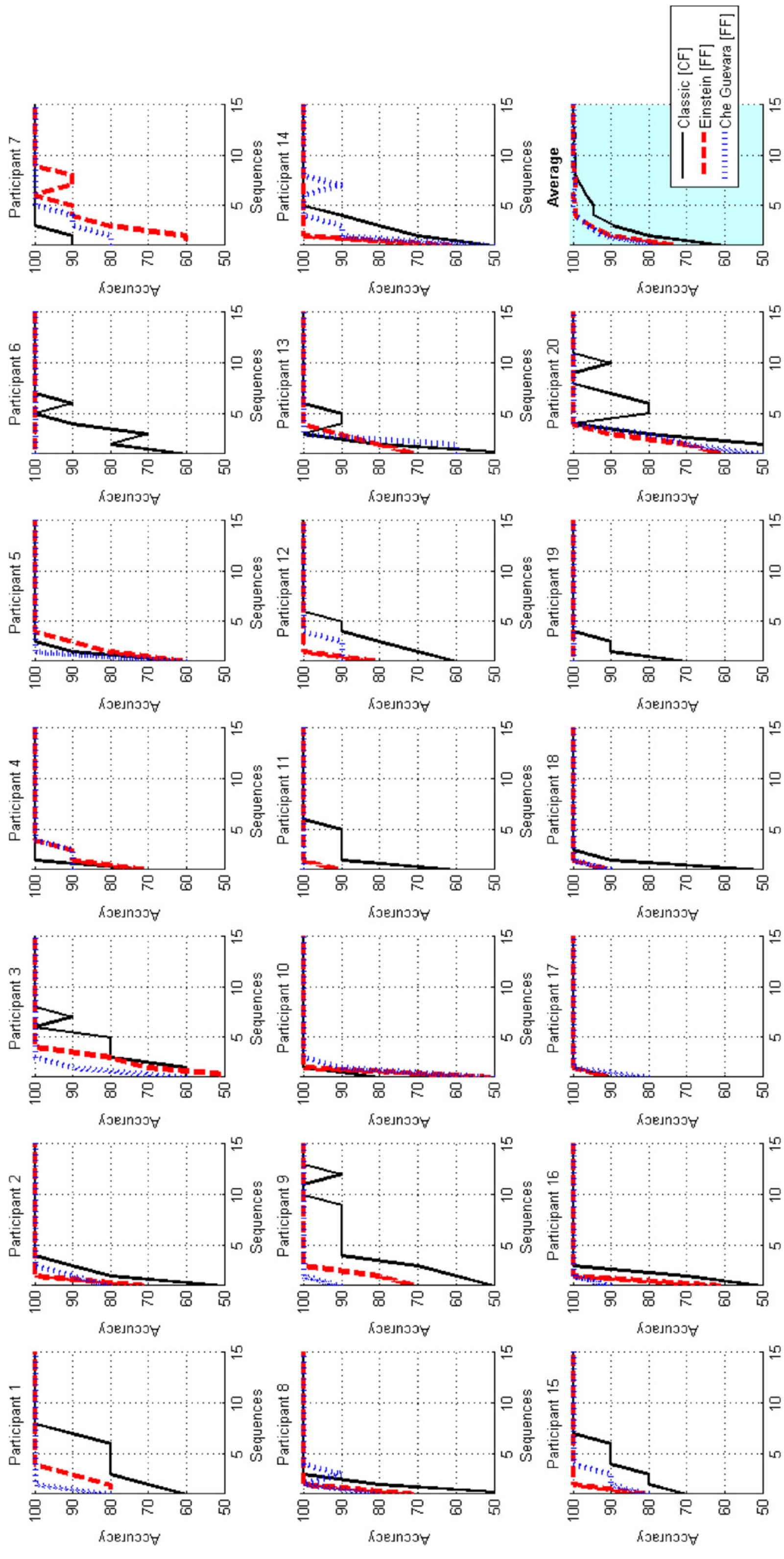


Figure 3. Classifier performance curves for all 15 stimulus intensification sequences. The three conditions ('classic "speller" (CF)', 'Einstein (FF)' and 'Che Guevara (FF)') are plotted for each participant. On the bottom right a grand average over participants is presented. As can be seen, for the majority of participants fewer sequences were necessary to achieve stable 70%, 90% and 100% accuracy levels in the FF conditions compared to CF.

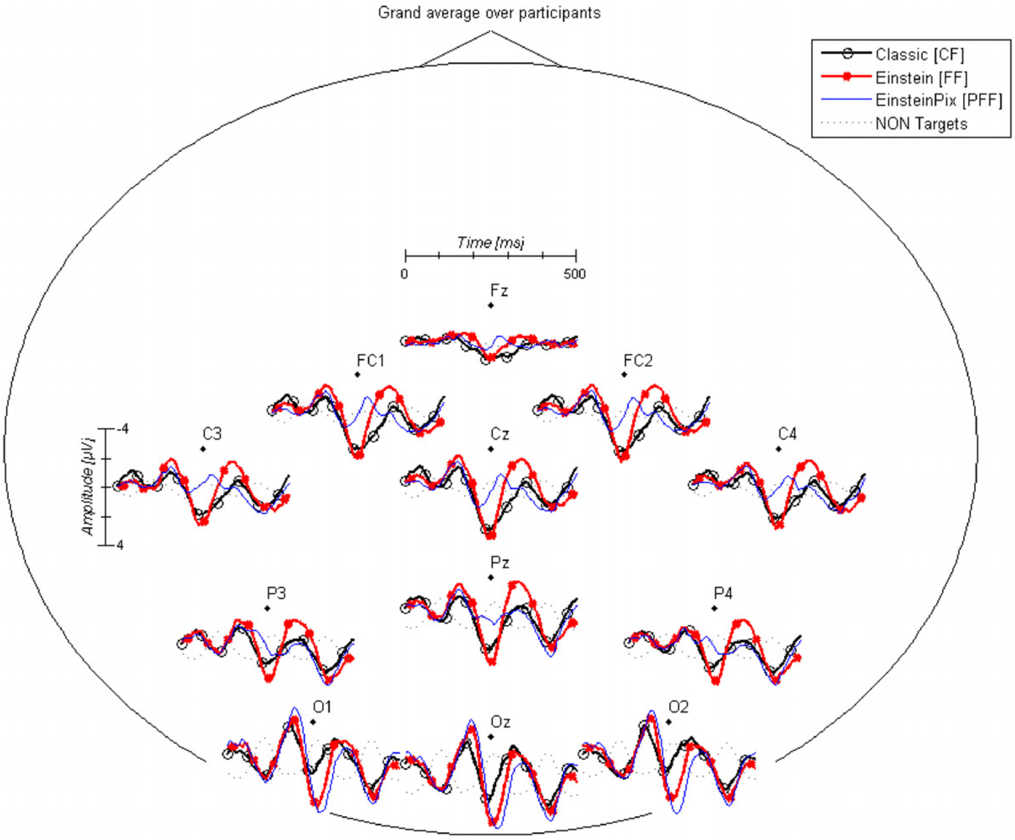


Figure 4. Topographic plot of all 12 electrodes for the three conditions CF, FF and PFF as a grand average over participants. In this case the FF and PFF conditions for the face of Einstein are presented. In the FF condition an N400f was elicited between 300 and 400 ms on centro-parietal electrode sites that was not present in any other condition.

high (>90%) accuracy level ($F(2, 38) = 11.7$, $p < 0.0001$, $\eta^2_{\text{partial}} = 0.38$). *Post-hoc* Tukey–Kramer comparisons revealed no significant difference between the two FF conditions for any of the accuracy levels. The difference between the CF condition and any of the FF conditions was highly significant (Tukey–Kramer: $p < 0.01$) for the 90% and 100% accuracy levels and significant (Tukey–Kramer: $p < 0.05$) for the 70% accuracy level. In sum, the presentation of flashing faces transparently superimposed on characters in the spelling matrix significantly decreased the number of sequences required as compared to simple highlighting of the characters.

3.2. Effects on ERPs

We further examined which ERPs contribute to this improvement in classification accuracy. Figure 4 shows a topographic plot of the event-related brain potential at different electrode sites averaged across all participants. The CF condition elicited a clear P300 with the largest amplitude at Cz. The same held true for the FF conditions (Einstein face in the example presented in figure 4). P300 amplitudes were not significantly different between CF and any of the FF conditions, but significantly decreased for the pixelated control of Einstein ($F(4, 76) = 9.34$, $p < 0.0001$, $\eta^2_{\text{partial}} = 0.32$). Presentation of familiar faces elicited two negative deflections, one between 130 and 200 ms (further referred to as N170 time window) and one between 300 and 400 ms (further referred to as N400f time window) with the maximum peak amplitudes at Pz. For further analysis the maximum peak amplitudes within these time windows were computed at Pz for all five conditions.

In the N170 time window mean peak amplitudes for both FF and PFF conditions were increased as compared to the CF condition ($F(4, 76) = 6.51$, $p < 0.0001$, $\eta^2_{\text{partial}} = 0.25$). In line with our expectations, this difference was highly significant for the Einstein FF condition (Tukey–Kramer: $p < 0.01$) and not significant for the corresponding PFF condition. However, the Che Guevara FF condition failed to reach significance and furthermore the corresponding PFF condition was highly significantly increased (Tukey–Kramer: $p < 0.01$).

For the N400f time window both FF conditions yielded highly significantly increased potentials as compared to CF and to both PFF conditions ($F(4, 76) = 24.17$, $p < 0.0001$, $\eta^2_{\text{partial}} = 0.56$), see figure 4. *Post-hoc* comparison confirmed high significance between CF and FF conditions

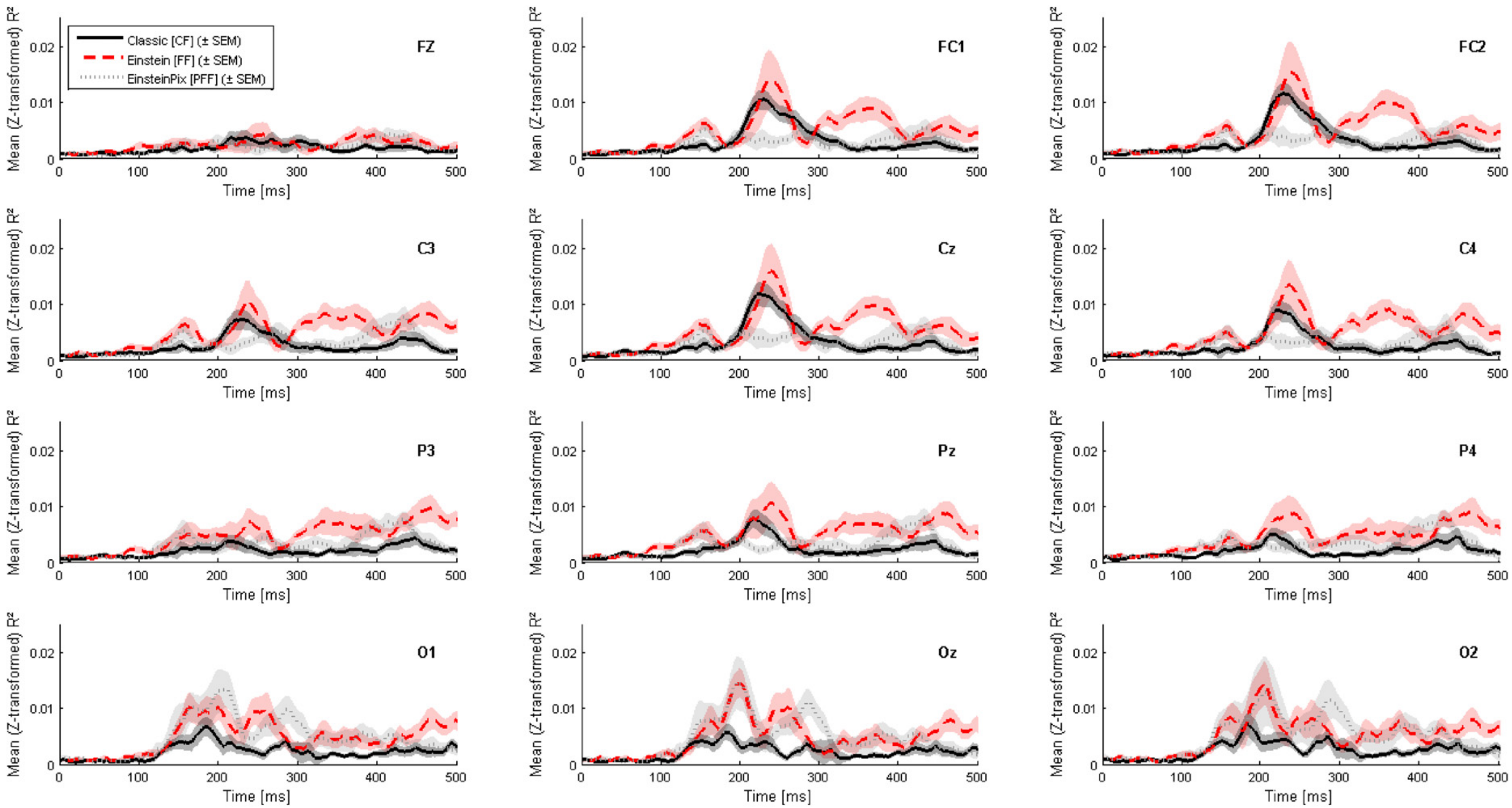


Figure 5. Mean determination coefficients over time for all electrode sites: values were Fisher-*Z* transformed and averaged. Here, the FF and PFF conditions for the face of Einstein are presented. As can be seen for FF and CF conditions the maximum determination was reached in the P300 time window. However, for the FF condition the N400f (i.e. Cz, 300–400 ms) further contributes significantly to determination.

(Tukey–Kramer: $p < 0.0001$). We further assessed this difference by comparing the maximum determination coefficients for the different conditions. Both FF conditions yielded significantly increased (highly significant in the case of Einstein, Tukey–Kramer: $p < 0.01$) maximum determination coefficients as compared to CF and PFF conditions ($F(4, 76) = 4.1$, $p < 0.01$, $\eta^2_{\mathrm{partial}} = 0.18$). As can be seen from figure 5 the maximum determination was found between 200 and 300 ms (the P300) for both FF and the CF condition as well. However the FF conditions (Einstein in figure 5) further yielded significantly increased determination in the N400f time window and a slightly increased determination in the N170 time window.

Taken together, ERP amplitudes and classification accuracy were significantly increased when famous faces were used for flashing as compared to the classic CF.

4. Discussion

In this study we compared famous faces (popular photographic representations of Albert Einstein and Che Guevara) superimposed transparently on the character matrix of the P300-Speller interface (FF) to the classic approach of simply flashing the characters (CF). In addition, we compared ERPs elicited in the FF condition to those elicited by stimuli with equal physical properties but no semantic or configuration information (PFF).

With such optimized stimulus presentation for eliciting ERPs, this study demonstrated increased speed of character selection in the widely used P300-Speller. In particular, the findings suggest that psychological salience of a stimulus can be exploited to elicit particularly strong ERPs not restricted to the P300. Consequently, 100% offline classification accuracy was achieved with significantly fewer sequences. Presentation of famous faces, rated as very familiar, produced (1) more ERPs that contributed to classification and (2) ERPs of larger amplitude which also contributed to better classification between attended versus unattended characters of the spelling matrix.

The significance of this approach becomes apparent when considering that the fluidity of communication by ‘typing’ clearly depends on the bit rate of character selection (number of correctly selected characters per time unit). Thus, achieving high bit rates is one of the main goals in BCI-based communication. The bit rate depends on both accuracy and speed of character selection. Speed is basically increased by decreasing ISIs (Sellers *et al* 2006a, 2006b) and decreasing the number of stimulus sequences used for averaging (Donchin *et al* 2000, Halder *et al* 2010, Kleih *et al* 2010). Sellers *et al* (2006a) achieved better classification accuracy and bit rate with small ISIs. Unfortunately, decreasing ISIs leads to decreased target-to-target intervals, which resulted in smaller P300 amplitudes and larger latencies (Gonsalvez and Polich 2002). As described in the introduction, Martens *et al* (2009) suggested an attempt to solve this problem. However, this approach is limited by the fact that to capture a certain ERP, a minimum number of data needs to be acquired and refractory periods cannot be completely eliminated. Thus, reducing the duration of single trials is inherently limited. Therefore, to further increase the speed of character selection one has to focus on reducing the number of stimulus sequences used for averaging. Unfortunately, this reduction inevitably decreases the signal-to-noise ratio and thus typically entails a drop in performance. As demonstrated in this study, using stimulus types that elicit particularly pronounced ERP responses serves well as a countermeasure to this issue, leading to an improved signal-to-noise ratio when reducing the number of sequences to average.

Our findings indicate that famous faces could reduce the overall time needed to spell a character on average by a factor of 1.8. Furthermore, at least twice as many of the participants (i.e. two-thirds of all participants) achieved a performance level of 70% already after the first stimulus sequence when using any of the FF paradigms as compared to the CF paradigm. This level has been described as the minimum level for communication (Kübler *et al* 2001b) and illustrates the potential of our new approach. Thus, in the majority of participants the FF paradigm would even allow for increasing the bit rate to maximum, that is flashing rows and columns only once.

For systematic further improvement of the conceptual idea behind the FF condition (i.e. exploiting semantic or psychological salience of stimuli to improve the ERP signal-to-noise ratio), it is important to better understand the reasons for a superior classification performance with well-known faces. Figures 4 and 5 indicate the main difference between all conditions in a time window of 300 to 400 ms at frontal, central and parietal electrodes with significantly higher amplitudes in the FF condition. This represents a face-specific N400f (Eimer 2000). The fact that the amplitudes in the PFF condition remain low supports this notion. Furthermore, there is an increased negative component between 150 and 200 ms in the FF condition. This component may well represent the expected N170. However, the PFF condition also showed increased negativity, which impedes a clear conclusion about this ERP component. Nevertheless, as this component does not appear as strong in the CF condition it constitutes a further signal difference that might allow for improved classification. Finally, the FF condition elicited a P300 at least of the same amplitude as the typical P300 classified in the classic CF paradigm (higher amplitude for Einstein). This effect is further underlined by a trend of higher determination coefficients in the P300 time window in the FF as compared to the CF condition (see figure 5). The fact that PFF conditions show decreased P300 amplitudes may be interpreted as an effect of smaller stimulus complexity and missing stimulus value, which has been shown to also influence P300 amplitudes (Johnson 1984, 1993).

4.1. Limitations and future experimentation

To further establish the significance of our findings, online application and studies with the target patient groups (e.g. those with ALS or other diseases leading to severe motor impairment) are mandatory. Since it has been shown that in a face recognition test ALS patients did perform equally well as age matched healthy controls (Abrahams *et al* 1996) we speculate that our results may well transfer to these patients. Although faces are typically recognized even when heavily degraded, a failure to focus eye-gaze may reduce effects in patients (e.g. Brunner *et al* (2010), Treder and Blankertz (2010)). Furthermore, covert attention to faces should be considered particularly when using the BCI in patient communication. Treder and Blankertz (2010) have suggested resolving issues of covert attention with a specialized character arrangement that clearly outperformed the classic matrix presentation in a covert attention condition. This approach could easily be adapted to our FF paradigm.

The generalizability of our results may be limited because the spelled words were rather short and did not comprise a representative set of characters from the complete alphabet. However, the stability of improvements across individual participants is promising with regard to generalization of effects to other samples.

Furthermore, as ERP latencies vary between participants, overlaps between them might occur. Hence, the N400f may partially cancel out the P300. The early peaking of the P300 in our data as well as reported in other BCI studies in general renders strong interaction relatively unlikely. Moreover, our findings indicate that the N400f contributed in combination with the P300 to classification accuracy in the FF conditions. Therefore, eliciting a N400f proved useful even if it would have cancelled out some P300 activity.

While there is ample evidence that the general P300-BCI paradigm is applicable for a wide range of individuals including patients, it is less clear whether all individuals will show reliably increased ERPs, in particular the N400f, in response to well-known faces. For example, our results indicate minor differences between the two FF conditions tested in this experiment. This may be attributable to physical properties of the stimuli. Thus a broader variety of face stimuli should be validated in terms of (1) familiarity, (2) brightness, (3) contrast and (4) chromatic differences. However, stimulus content may also influence the ERPs that are the semantic or psychological properties of the stimuli. Thus, stimuli varying in terms of individual relevance with regard to psychological (e.g. valence, arousal) or semantic properties (e.g. individual attraction) should be examined. For example, a recent study demonstrated that celebrities elicit a stronger N400f than unknown faces, but the strongest ERPs were elicited by pictures of family members (Touryan *et al* 2011). This could not only serve the goal of further improving classification but might as well increase patient acceptance of the BCI because it can be performed with personally meaningful and liked pictures. Importantly, implementing personally meaningful stimuli in the P300-BCI is straightforward and feasible, and could thus be provided for all BCI users, allowing an individual to achieve high classification at the maximum bit rate.

5. Conclusion

This study demonstrated that superimposing characters of the P300-Speller with famous faces significantly increases the bit rate. The results from $N = 20$ participants indicated a highly significant improvement of classification rates, primarily based on ERP components that are typically elevated when familiar faces are processed, in particular the N400f. This may have a clinically significant impact by increasing communication speed and accuracy with the P300-Speller in patients with severe motor impairment.

Acknowledgments

This work is supported by the European ICT Programme Projects FP7-224631. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein. SMS had the idea of using flashing faces in the ERP-BCI; SMS, TK and AK designed the study; TK programmed the speller paradigm; CG and TK collected the data; TK and SMS analyzed the data; TK, SMS, CG and AK discussed the results; TK drafted the manuscript, AK, SMS and CG critically revised the manuscript. All gave their final approval to the version to be published.

References

Abrahams S, Goldstein L H, Kew J J, Brooks D J, Lloyd C M, Frith C D and Leigh P N 1996 Frontal lobe dysfunction in amyotrophic lateral sclerosis. A PET study *Brain* **119** (Pt 6) 2105–20

Allison B Z and Pineda J A 2003 ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system *IEEE Trans. Neural Syst. Rehabil. Eng.* **11** 110–3

Allison B Z and Pineda J A 2006 Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system *Int. J. Psychophysiol.* **59** 127–40

Bentin S, Allison T, Puce A, Perez E and McCarthy G 1996 Electrophysiological studies of face perception in humans *J. Cogn. Neurosci.* **8** 551–65

Bianchi L, Sami S, Hillebrand A, Fawcett I P, Quitadamo L R and Seri S 2010 Which physiological components are more suitable for visual ERP based brain–computer interface? A preliminary MEG/EEG study *Brain Topogr.* **23** 180–5

Birbaumer N, Ghanayim N, Hinterberger T, Iversen I, Kotchoubey B, Kübler A, Perelmouter J, Taub E and Flor H 1999 A spelling device for the paralysed *Nature* **398** 297–8

Blankertz B, Lemm S, Treder M, Haufe S and Müller K-R 2011 Single-trial analysis and classification of ERP components—a tutorial *NeuroImage* **56** 814–25

Brunner P, Joshi S, Briskin S, Wolpaw J R, Bischof H and Schalk G 2010 Does the 'P300' speller depend on eye gaze? *J. Neural Eng.* **7** 056013

Donchin E, Spencer K M and Wijesinghe R 2000 The mental prosthesis: assessing the speed of a P300-based brain-computer interface *IEEE Transactions on Rehabilitation Engineering: a Publication of the IEEE Engineering in Medicine and Biology Society* **8** 174–9

Eimer M 2000 Event-related brain potentials distinguish processing stages involved in face perception and recognition *Clin. Neurophysiol.* **111** 694–705

Farwell L A and Donchin E 1988 Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials *Electroencephalogr. Clin. Neurophysiol.* **70** 510–23

Frye G E, Hauser C K, Townsend G and Sellers E W 2011 Suppressing flashes of items surrounding targets during calibration of a P-300-based brain–computer interface improves performance *J. Neural Eng.* **8** 025024

Gonsalvez C L and Polich J 2002 P300 amplitude is determined by target-to-target interval *Psychophysiology* **39** 388–96

Gratton G, Coles M G and Donchin E 1983 A new method for off-line removal of ocular artifact *Electroencephalogr. Clin. Neurophysiol.* **55** 468–84

Guger C, Daban S, Sellers E, Holzner C, Krausz G, Carabalona R, Gramatica F and Edlinger G 2009 How many people are able to control a P300-based brain–computer interface (BCI)? *Neurosci. Lett.* **462** 94–8

Halder S, Rea M, Andreoni R, Nijboer F, Hammer E M, Kleih S C, Birbaumer N and Kübler A 2010 An auditory oddball brain–computer interface for binary choices *Clin. Neurophysiol.* **121** 516–23

Hanagasi H A, Gurvit I H, Ermutlu N, Kaptanoglu G, Karamursel S, Idrisoglu H A, Emre M and Demiralp T 2002 Cognitive impairment in amyotrophic lateral sclerosis: evidence from neuropsychological investigation and event-related potentials *Brain Res. Cogn. Brain Res.* **14** 234–44

Henson R N, Goshen-Gottstein Y, Ganel T, Otten L J, Quayle A and Rugg M D 2003 Electrophysiological and haemodynamic correlates of face perception, recognition and priming *Cereb. Cortex* **13** 793–805

Jin J, Horki P, Brunner C, Wang X, Neuper C and Pfurtscheller G 2010 A new P300 stimulus presentation pattern for EEG-based spelling systems *Biomed. Tech. Biomed. Eng.* **55** 203–10

Johnson R Jr 1984 P300: a model of the variables controlling its amplitude *Ann. New York Acad. Sci.* **425** 223–9

Johnson R Jr 1993 On the neural generators of the P300 component of the event-related potential *Psychophysiology* **30** 90–7

Kleih S C *et al* 2011 Out of the frying pan into the fire—the P300 based BCI faces real world challenges *Prog. Brain Res.* **194** at press

Kleih S C, Nijboer F, Halder S and Kübler A 2010 Motivation modulates the P300 amplitude during brain–computer interface use *Clin. Neurophysiol.* **121** 1023–31

Krusienski D, Sellers E, Cabestaing F, Bayoudh S, McFarland D, Vaughan T and Wolpaw J 2006 A comparison of classification techniques for the P300 Speller *J. Neural Eng.* **3** 299–305

Krusienski D J, Sellers E W, McFarland D J, Vaughan T M and Wolpaw J R 2008 Toward enhanced P300 speller performance *J. Neurosci. Methods* **167** 15–21

Kübler A and Birbaumer N 2008 Brain–computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? *Clin. Neurophysiol.* **119** 2658–66

Kübler A, Kotchoubey B, Kaiser J, Wolpaw J R and Birbaumer N 2001a Brain–computer communication: unlocking the locked in *Psychol. Bull.* **127** 358–75

Kübler A, Neumann N, Kaiser J, Kotchoubey B, Hinterberger T and Birbaumer N P 2001b Brain–computer communication: self-regulation of slow cortical potentials for verbal communication *Arch. Phys. Med. Rehabil.* **82** 1533–9

Lotte F, Congedo M, Lécuyer A, Lamarche F and Arnaldi B 2007 A review of classification algorithms for EEG-based brain–computer interfaces *J. Neural Eng.* **4** R1–13

Martens S M M, Hill N J, Farquhar J and Schölkopf B 2009 Overlap and refractory effects in a brain–computer interface speller based on the visual P300 event-related potential *J. Neural Eng.* **6** 026003

McFarland D J, Sarnacki W A, Townsend G, Vaughan T and Wolpaw J R 2011 The P300-based brain–computer interface (BCI): effects of stimulus rate *Clin. Neurophysiol.* **122** 731–7

Nijboer F *et al* 2008 A P300-based brain–computer interface for people with amyotrophic lateral sclerosis *Clin. Neurophysiol.* **119** 1909–16

Picton T W 1992 The P300 wave of the human event-related potential *J. Clin. Neurophysiol.* **9** 456–79

Polich J 2007 Updating P300: an integrative theory of P3a and P3b *Clin. Neurophysiol.* **118** 2128–48

Polich J, Alexander J E, Bauer L O, Kuperman S, Morzorati S, O'Connor S J, Porjesz B, Rohrbaugh J and Begleiter H 1997 P300 topography of amplitude/latency correlations *Brain Topogr.* **9** 275–82

Salvaris M and Sepulveda F 2009 Visual modifications on the P300 speller BCI paradigm *J. Neural Eng.* **6** 046011

Schalk G, McFarland D J, Hinterberger T, Birbaumer N and Wolpaw J R 2004 BCI2000: a general-purpose brain–computer interface (BCI) system *IEEE Trans. Biomed. Eng.* **51** 1034–43

Sellers E W, Krusienski D J, McFarland D J, Vaughan T M and Wolpaw J R 2006a A P300 event-related potential brain–computer interface (BCI): the effects of matrix size and inter stimulus interval on performance *Biol. Psychol.* **73** 242–52

Sellers E W, Kübler A and Donchin E 2006b Brain–computer interface research at the University of South Florida Cognitive Psychophysiology Laboratory: the P300 Speller *IEEE Trans. Neural Syst. Rehabil. Eng.* **14** 221–4

Sellers E W, Vaughan T M and Wolpaw J R 2010 A brain–computer interface for long-term independent home use *Amyotrophic Lateral Scler.* **11** 449–55

Sutton S, Braren M, Zubin J and John E R 1965 Evoked-potential correlates of stimulus uncertainty *Science* **150** 1187–8

Takano K, Komatsu T, Hata N, Nakajima Y and Kansaku K 2009 Visual stimuli for the P300 brain–computer interface: a comparison of white/gray and green/blue flicker matrices *Clin. Neurophysiol.* **120** 1562–6

Touryan J, Gibson L, Horne J H and Weber P 2011 Real-time measurement of face recognition in rapid serial visual presentation *Front. Psychol.* **2** 42

Townsend G *et al* 2010 A novel P300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns *Clin. Neurophysiol.* **121** 1109–20

Treder M S and Blankertz B 2010 (C)overt attention and visual speller design in an ERP-based brain–computer interface *Behav. Brain Funct.* **6** 28

2.5 Kaufmann, T., Schulz, S. M., Köblitz, A., Renner, G., Wessig, C., and Kübler, A. (2013). Face stimuli effectively prevent brain-computer interface inefficiency in patients with neurodegenerative disease. Clinical Neurophysiology

Abstract: *Objectives*: Recently, we proposed a new stimulation paradigm for brain computer interfaces (BCI) based on event-related potentials (ERP), i.e. flashing characters with superimposed pictures of well-known faces. This new face flashing (FF) paradigm significantly outperformed the commonly used character flashing (CF) approach, i.e. simply highlighting characters. *Methods*: In the current study we assessed the impact of face stimuli on BCI inefficiency in patients with neurodegenerative disease, i.e. on their inability to communicate by means of a BCI. Healthy participants (N = 16) and those with neurodegenerative disease (N = 9) performed spelling tasks using CF and FF paradigms. *Results*: Online performance with FF was significantly increased as compared to CF in both, healthy and impaired users. Importantly, two patients who were classified ”highly inefficient” with the classic CF stimulation were able to spell with high accuracy using FF. Our results particularly emphasize great benefit of the FF paradigm for those users displaying low signal-to-noise ratio of the recorded ERPs in the classic stimulation approach. *Conclusion*: In conclusion, we confirm previously reported results now systematically validated in an online setting and display specifically beneficial effects of FF for motor-impaired users. *Significance*: The FF paradigm thus constitutes a big step forward against the BCI inefficiency phenomenon.

Link to publication source: Clinical Neurophysiology, Elsevier B.V.

Clinical Neurophysiology 124 (2013) 893–900

Contents lists available at SciVerse ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Face stimuli effectively prevent brain–computer interface inefficiency in patients with neurodegenerative disease

Tobias Kaufmann [a,*], Stefan M. Schulz [a,b], Anja Köblitz [a], Gregor Renner [c], Carsten Wessig [d], Andrea Kübler [a,e]

[a] Institute of Psychology, University of Würzburg, Marcusstr. 9-11, 97070 Würzburg, Germany
[b] Comprehensive Heart Failure Center (CHFC), University of Würzburg, Straubmühlweg 2a, 97078 Würzburg, Germany
[c] Catholic University of Applied Sciences, Karlstr. 63, 79104 Freiburg, Germany
[d] Department of Neurology, University of Würzburg, Josef-Schneider-Str. 11, 97080 Würzburg, Germany
[e] Institute of Medical Psychology and Behavioural Neurobiology, University of Tübingen, Gartenstr. 29, 72074 Tübingen, Germany

See Editorial, pages 831–833

A R T I C L E I N F O

Article history:
Accepted 7 November 2012
Available online 14 December 2012

Keywords:
Brain computer interface
Event related potentials
Neurodegenerative disease
Communication
BCI inefficiency
P300-speller
Online classification accuracy
Bit rate

H I G H L I G H T S

- This study shows how to effectively overcome brain computer interface (BCI) inefficiency in patients with neurodegenerative disease.
- Online performance was significantly increased in healthy participants ($N = 16$) and those with neurodegenerative disease ($N = 9$) when using faces as stimulus material in an ERP–BCI paradigm.
- Importantly, two patients who were highly inefficient with the classic BCI paradigm spelled at high accuracy levels with the face flashing paradigm.

A B S T R A C T

Objectives: Recently, we proposed a new stimulation paradigm for brain computer interfaces (BCI) based on event-related potentials (ERP), i.e. flashing characters with superimposed pictures of well-known faces. This new face flashing (FF) paradigm significantly outperformed the commonly used character flashing (CF) approach, i.e. simply highlighting characters.
Methods: In the current study we assessed the impact of face stimuli on BCI inefficiency in patients with neurodegenerative disease, i.e. on their inability to communicate by means of a BCI. Healthy participants ($N = 16$) and those with neurodegenerative disease ($N = 9$) performed spelling tasks using CF and FF paradigms.
Results: Online performance with FF was significantly increased as compared to CF in both, healthy and impaired users. Importantly, two patients who were classified "highly inefficient" with the classic CF stimulation were able to spell with high accuracy using FF. Our results particularly emphasize great benefit of the FF paradigm for those users displaying low signal-to-noise ratio of the recorded ERPs in the classic stimulation approach.
Conclusion: In conclusion, we confirm previously reported results now systematically validated in an online setting and display specifically beneficial effects of FF for motor-impaired users.
Significance: The FF paradigm thus constitutes a big step forward against the BCI inefficiency phenomenon.

1. Introduction

Brain computer interfaces (BCI) based on event-related potentials (ERP) provide a communication channel independent from muscular control, thus, potentially suited for patients with neurodegenerative diseases or severe motor impairment due to other causes such as brainstem stroke (Farwell and Donchin, 1988; Sellers and Donchin, 2006; Nijboer et al., 2008; for review Kleih et al., 2011; Mak et al., 2011). Such ERP–BCIs utilize a so-called oddball paradigm, i.e. presenting a rare target stimulus within a set of irrelevant stimuli. Commonly, users are presented with a matrix consisting of characters that are highlighted (flashed) in random order (see Fig. 1A). Communication is established by focusing attention on the intended character and counting the number of flashes. The attended target stimuli elicit

* Corresponding author. Address: Department of Psychology I, Section Intervention Psychology; Marcusstr. 9-11, 97070 Würzburg, Germany. Tel.: +49 931 31 81084; fax: +49 931 31 82424.
E-mail address: tobias.kaufmann@uni-wuerzburg.de (T. Kaufmann).

1388-2457/$36.00 © 2012 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.clinph.2012.11.006

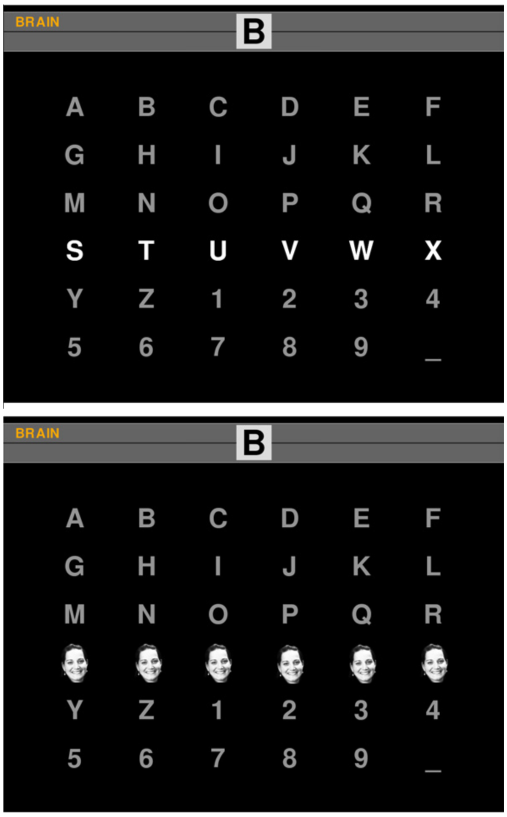


Fig. 1. (A) Classic character flashing paradigm [CF] in which rows and columns are highlighted randomly. Users focus their attention on the target stimulus and count the number of target flashes. (B) BCI paradigm proposed by Kaufmann et al. (2011) in which characters are overlaid with faces and users count the number of face flashes [FF]. In this figure, the face stimulus displays a personally known face. Besides this stimulus type, a famous photograph of Albert Einstein and an unfamiliar face were used.

ERPs (for review, Polich, 2007) in the electroencephalogram (EEG) that can be classified and hence indicate the intended character for communication (Farwell and Donchin, 1988).

Although being fast and reliable in healthy participants, studies on ERP–BCI use by impaired participants reveal large inter-individual variations in achieved BCI performance (Nijboer et al., 2008; Kübler and Birbaumer, 2008; for review, Mak et al., 2011). Reliable communication does not require perfect spelling performance since human communication partners or advanced predictive text algorithms are able to correct for spelling errors and extract informational content. Yet, BCI performance of patients is often even below a minimum level of accuracy required for basic communication (e.g., 70% accuracy as suggested by Kübler et al., 2001). The term "BCI illiteracy" has often been used to describe non-successful BCI use (e.g. Kübler and Müller, 2007; Vidaurre and Blankertz, 2010; Blankertz et al., 2010), but was suggested to be replaced by BCI inefficiency to better stress that the inability may be inherent in the system, not in the user (Kübler et al., 2011). The major goal of ERP–BCI research can thus be described as a 2-step process, (1) establishing a sufficient accuracy level for communication and if successful (2) increasing spelling speed without decreasing accuracy, i.e. increasing communication bit rate (correctly spelled characters per time unit).

Recently we addressed this issue by developing a new paradigm which aims at increasing the signal-to-noise ratio (SNR) within the entire classification time window by eliciting additional target specific ERPs (Kaufmann et al., 2011, see Fig. 1B). When flashing characters with transparently superimposed well-known faces, ERPs involved in face processing are elicited (N170, N400f) and were found to significantly increase classification accuracy in an offline setting. For example when well-known faces were used as stimulus material (face flashing, FF) a stable level of 100% offline accuracy was achieved in healthy participants with significantly fewer sequences necessary for classification than with classic character flashing (character flashing, CF).

Consequently with FF, it was possible to achieve high accuracy levels in an online setting using single trial classification in seven healthy participants (Zhang et al., 2012). However, until now it has not been known how the FF paradigm affects BCI inefficiency in severely motor-impaired end-users. Thus, this study systematically validated online classification accuracy in both healthy and motor-impaired BCI users.

Furthermore, our study explored optimization of stimulus material. Jin and colleagues (2012) investigated if face emotion and/or motion of FF stimuli may increase spelling accuracy, yet no difference was found between these stimuli. Herein we investigated the role of face familiarity. Initially we proposed use of famous faces (Kaufmann et al., 2011). Touryan and colleagues (2011) found that faces of family members elicited larger N400 potentials than celebrity faces. Such personally known faces may also increase end-user acceptance of the stimulus material due to personal meanings of the pictures (Kaufmann et al., 2011). Jin and colleagues (2012) used one face that was personally known by all participants (fellow student) whereas Zhang and colleagues (2012) used a prior unfamiliar face. In the current study, we compared these conditions (unfamiliar FF, famous FF, and personally known FF) to investigate effects of face familiarity on spelling accuracy.

2. Methods

2.1. Stimulus material

In our previous study (Kaufmann et al., 2011) we used two face flashing (FF) stimuli, i.e. famous photographs of Albert Einstein and Ernesto 'Che' Guevara. As we found no difference in classification accuracy between these FF conditions, we herein used the famous face that was rated as marginally more familiar, i.e. the face of Albert Einstein. Besides this famous face we included a prior unfamiliar face and personally known faces for comparison across FF stimuli. All participants sent a photograph of a personally known face (family member or close friend) prior to the experiment, so the pictures could be edited for use as stimulus material. All FF stimuli were comparable in displaying a face on a black background. Apart from FF stimuli, one condition comprised the commonly used character flashing approach (CF), i.e. simply highlighting the matrix characters.

2.2. Participants

Participants were $N = 16$ healthy BCI novices (11 women; mean age 23.69 years, $SD = 2.6$, range 19–33) and $N = 9$ patients with neurodegenerative diseases (eight men; mean age 50.00 years, $SD = 15.21$, range 26–72). Table 1 provides detailed overview on medical diagnosis and physical state of individual patients. The experiment was approved by the local Ethical Review Board (University of Würzburg, Germany) and conducted in accordance with standard ethical guidelines as defined by the Declaration of

Table 1
Description of patient sample.

	Age	Gender	Diagnosis	Date of diagnosis	Description of current state
Patient 1	27	Male	SMA, type II	1986	Wheelchair bound. Severely paralyzed except for residual finger movements (used for wheelchair control) and control of facial muscles. Intact speech
Patient 2	25	Male	SMA, type II	1987	Wheelchair bound. Severely paralyzed except for residual finger movements (used for wheelchair control) and control of facial muscles. Intact speech
Patient 3	48	Male	ALS	2007	Limited in arm and foot movements. Able to walk on short distances using foot orthosis. Intact speech
Patient 4	59	Male	ALS	2007	Wheelchair bound. Artificially ventilated. Intact speech
Patient 5	53	Male	SBMA	2004	Usually wheelchair bound but able to walk on short distances. Intact speech
Patient 6	72	Male	MD	1985	Able to walk. Requires regular head rest. Intact speech
Patient 7	59	Male	ALS, bulbar	2011	Able to walk. Severely impaired speech
Patient 8	56	Male	SBMA	2003	Able to walk. Occasional impairment in movement and speech
Patient 9	51	Female	ALS	2003	Severely paralyzed except for residual movement of one finger and control of facial muscles. Severely impaired speech (understandable only for a set of acquaintances through reading lips)

SMA, spinal muscular atrophy; ALS, amyotrophic lateral sclerosis; SBMA, spinobulbar muscular atrophy; Kennedy's syndrome; MD, muscular dystrophy.

Helsinki (World Medical Association) and the European Council's Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine (Convention on Human Rights and Biomedicine). All participants signed informed consent prior to the study.

2.3. Experimental design

All healthy participants spelled words with all four stimulus conditions (3 FF: famous face, personally known face, unfamiliar face; 1 CF: classic character flashing; see Section 2.1). Participants from the patient sample spelled words with three conditions (2 FF: famous face, personally known face; 1 CF: classic character flashing). We omitted the unfamiliar face condition for this group to shorten the experimental schedule and thereby reduce burden for the patients. The order of the conditions was randomized across participants. Prior to online spelling, one calibration session per stimulus condition was performed, i.e. participants spelled the word 'BRAINPOWER' for all stimulus conditions separately with 15 sequences per character (one sequence comprised each row and column flashed once). Subsequent to the four calibration sessions, classification was performed and classifier weights adjusted for each stimulus condition (stepwise-linear discriminant analysis, 800 ms post-stimulus; see e.g. Krusienski et al., 2008).

In the online sessions (OS; one session comprised 1 run per paradigm) we gradually decreased the number of sequences (NoS) used for flashing the matrix in 5 steps (OS_1: NoS = 10; OS_2: NoS = 6; OS_3: NoS = 3; OS_4: NoS = 2; OS_5: NoS = 1). This study design enabled to systematically validate the impact of stimulus material on BCI performance in that decreasing the number of sequences entails higher probability of miss-classification. In all online sessions, participants were required to spell the word 'BRAIN' to allow for direct comparison of user performance between sessions.

2.4. Equipment, data acquisition and feedback presentation

EEG was obtained from 12 passive Ag/AgCl electrodes at positions Fz, FC1, FC2, C3, Cz, C4, P3, Pz, P4, O1, Oz, O2 with mastoid ground and reference. Signals were amplified using a g.USBamp amplifier (*g.tec Medical Engineering GmbH*, *Austria*) and recorded at 512 Hz sampling rate using the BCI2000 software (Schalk et al., 2004).

Spelling paradigms were implemented in Python 2.5 and connected to BCI2000 via user datagram protocol (UDP). In a visually displayed character matrix of size 6×6 (screen size: 19″; refresh rate: 60 Hz; size of matrix: 30×20 cm^2; distance of participants to screen: ~50 cm), target rows and columns were either highlighted (CF) or overlaid with faces (FF, see Section 2.1). Stimulus timing was set to 33.3 ms stimulus duration and 133.3 ms inter stimulus interval. During the online sessions, the selected character was presented on top of the character matrix after each sequence, thus providing immediate feedback to the participants.

2.5. ERP analysis

EEG data was filtered with 0.1 Hz high and 30 Hz low pass. For ERP analysis, we deleted overlapping trials (e.g. double flashes) from the data set. The remaining data showed no overlap for at least 466 ms post-stimulus (for details on the procedure see Kaufmann et al., 2011). According to the results obtained from the grand average ERP we defined ERP amplitudes as follows: N170 (minimum voltage between 120 and 200 ms), P300 (maximum voltage between 200 and 300 ms) and N400f (minimum voltage between 280 and 400 ms). Data analysis was performed in Matlab 2011b (The Mathworks, USA).

2.6. Data analysis

Online spelling accuracy was expressed as the percentage of correctly typed characters during the online sessions. Non-parametric procedures were used for statistical data analysis because data was not normally distributed. For each NoS (see Section 2.3), the achieved online spelling accuracy was compared across the different stimulus conditions using Kruskal–Wallis tests. Post-hoc analysis was performed with Mann–Whitney-*U* tests. Bonferroni correction to alpha levels is indicated in the results section. Statistical analysis was performed in SPSS 20 (IBM Corporation, USA).

Bit rate was computed as described by Wolpaw and colleagues (1998); derived from Shannon and Weaver, 1964; Pierce, 1980) and is expressed in communicated bits/min. This measure includes duration of flash, inter-stimulus and inter-sequence interval.

3. Results

3.1. Spelling performance

All participants completed all scheduled sessions, except for patients 1 and 2, who skipped the last online session due to strain (OS_5, i.e. session with NoS = 1 for all three stimulus conditions). Classification accuracy estimated offline from calibration data was in line with our previous report (Kaufmann et al., 2011) in that face stimuli effectively increased offline classification accuracy (Fig. 2, CF vs. FF).

The superior efficiency of the FF paradigm emerged particularly when exposing participants to the online setting with adjusted numbers of sequences (Fig. 3). Performance was significantly different between stimulus conditions in healthy participants across online sessions with 1–6 sequences (OS_2 to OS_5 significantly different; all $H(3) > 16.47$, all $p \leqslant .001$; Bonferroni adjusted alpha level: $\alpha = .0083$) and across all online sessions in patients (all $H(2) > 8.85$, all $p \leqslant .012$; Bonferroni adjusted alpha level: $\alpha = .0167$). Post-hoc comparison of online performance revealed no significant difference between any of the FF stimulus conditions but significant differences between CF and all FF conditions (for patients in OS_2 to OS_5; for healthy participants in OS_2 to OS_5 except for CF vs. unfamiliar face in OS_3 and CF vs. personally known face in OS_4). Tables 2 and 3 provide full statistics.

The practical relevance of the strong benefit achieved with the FF stimuli becomes apparent in the patient sample (see Fig. 3A and B). With the classic CF paradigm, their online performance considerably decreased when reducing the number of sequences. None of the patients performed better than 60% accuracy in OS_4 (average: 28.9%) and none better than 40% in OS_5 (average: 11.4%). In the latter session, 4 of 7 patients even miss-spelled all letters using CF. However, such BCI inefficiency was not visible when exposed to the FF paradigms. When flashing the matrix with one sequence only, performance remained at least 60% in all patients (average: 85.71% with famous FF; 77.14% with personally known face FF). This remaining performance level allowed for spelling at high bit rates (Fig. 4).

As usual, significant differences between healthy and motor impaired participants were found for the performance achieved with CF ($Z = -2.13$, $p < .017$; Bonferroni adjusted alpha level: $\alpha = .0167$;

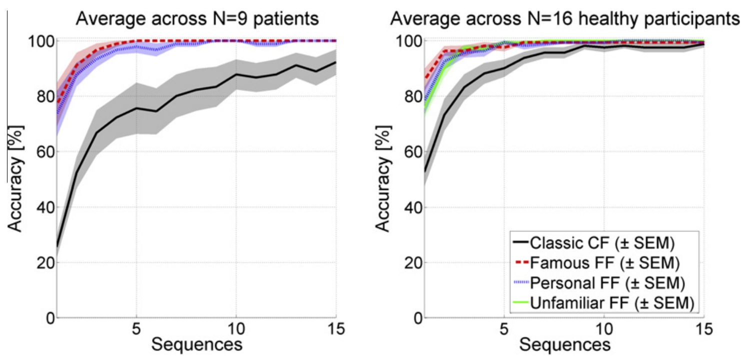


Fig. 2. Average classification accuracy and standard errors of the mean (SEM) estimated offline from calibration data of $N = 16$ healthy participants and $N = 9$ patients with neurodegenerative disease.

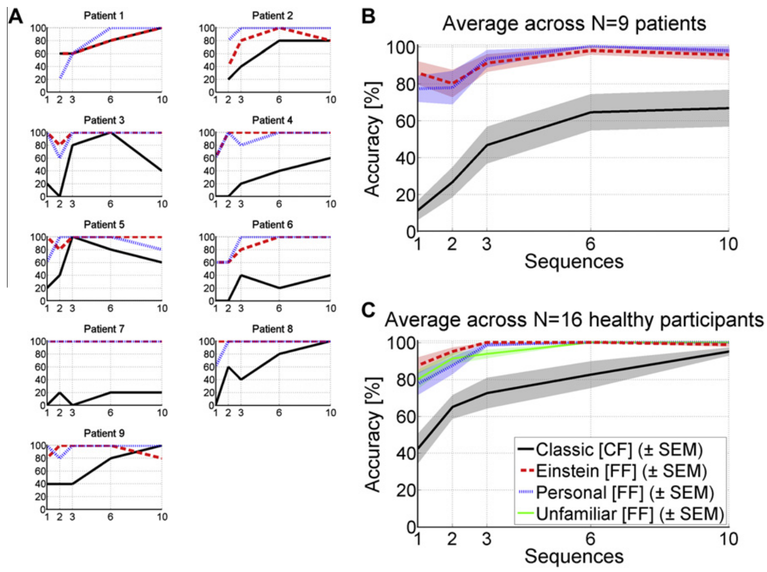


Fig. 3. Performance in online sessions (OS) from $N = 16$ healthy participants and $N = 9$ patients with neurodegenerative disease. The number of sequences used for flashing the matrix was systematically reduced in 5 steps from 10 to 1 to estimate the impact of FF stimuli on bit rate (see Section 2.3, OS_1: NoS = 10; OS_2: NoS = 6; OS_3: NoS = 3; OS_4: NoS = 2; OS_5: NoS = 1). (A) Single patient performance for all five online sessions. (B) Average over patients and (C) healthy participants including standard errors of the mean (SEM).

Table 2
Post-hoc Mann–Whitney-*U* tests for direct comparison across all four stimulus conditions of performance achieved by $N = 16$ healthy participants.

		Famous FF	Unfamiliar FF	Personally known FF
NoS = 1 (OS_5)	Classic CF	$Z = -3.83$ $p < .0001^*$	$Z = -3.23$ $p < .0012^*$	$Z = -3.02$ $p = .0025^*$
	Famous FF		$Z = -1.14$ $p = .25$	$Z = -1.27$ $p = .20$
	Unfamiliar FF			$Z = -0.20$ $p = .84$
NoS = 2 (OS_4)	Classic CF	$Z = -3.55$ $p < .0004^*$	$Z = -3.04$ $p < .0023^*$	$Z = -2.61$ $p = .0091$
	Famous FF		$Z = -0.84$ $p = .40$	$Z = -1.02$ $p = .31$
	Unfamiliar FF			$Z = -0.29$ $p = .78$
NoS = 3 (OS_3)	Classic CF	$Z = -3.17$ $p < .0015^*$	$Z = -1.77$ $p = .08$	$Z = -2.83$ $p < .0047^*$
	Famous FF		$Z = -2.40$ $p = .017$	$Z = -1.00$ $p = .32$
	Unfamiliar FF			$Z = -1.78$ $p = .08$
NoS = 6 (OS_2)	Classic CF	$Z = -2.66$ $p < .0078^*$	$Z = -2.66$ $p < .0078^*$	$Z = -2.66$ $p < .0078^*$
	Famous FF		$Z = \text{inf}$ $p > .99$	$Z = \text{inf}$ $p > .99$
	Unfamiliar FF			$Z = \text{inf}$ $p > .99$
NoS = 10 (OS_1)	Classic CF	$Z = -1.43$ $p = .15$	$Z = -2.10$ $p = .04$	$Z = -2.10$ $p = .03$
	Famous FF		$Z = -1.00$ $p = .32$	$Z = -1.00$ $p = .32$
	Unfamiliar FF			$Z = \text{inf}$ $p > .99$

No significant difference was found between any of the FF conditions, but large differences between FF and CF. NoS, number of sequences; CF, character flashing; FF, face flashing.
* Significant at Bonferroni adjusted alpha levels of $p \leqslant .0083$.

Table 3
Post-hoc Mann–Whitney-*U* tests for direct comparison across all three stimulus conditions of performance achieved by $N = 9$ patients.

		Famous FF	Personally known FF
NoS = 1 (OS_5)	Classic CF	$Z = -3.21$ $p < .0013^*$	$Z = -3.22$ $p < .0013^*$
	Famous FF		$Z = -0.79$ $p = .432$
NoS = 2 (OS_4)	Classic CF	$Z = -3.18$ $p < .0014^*$	$Z = -3.00$ $p < .0027^*$
	Famous FF		$Z = -0.05$ $p = .963$
NoS = 3 (OS_3)	Classic CF	$Z = -2.94$ $p < .0033^*$	$Z = -3.07$ $p < .0021^*$
	Famous FF		$Z = -0.45$ $p = .654$
NoS = 6 (OS_2)	Classic CF	$Z = -3.18$ $p < .0015^*$	$Z = -3.54$ $p < .0004^*$
	Famous FF		$Z = -1.0$ $p = .317$
NoS = 10 (OS_1)	Classic CF	$Z = -2.24$ $p = .025$	$Z = -2.52$ $p < .012^*$
	Famous FF		$Z = -0.62$ $p = .539$

No significant difference was found between any of the FF conditions, but large differences between FF and CF. NoS, number of sequences; CF, character flashing; FF, face flashing.
* Significant at Bonferroni adjusted alpha level of $p \leqslant .0167$).

comparison of single-sequence performance), i.e. patients performed considerably worse than healthy users. However, this group difference vanished in the FF conditions and patients per-

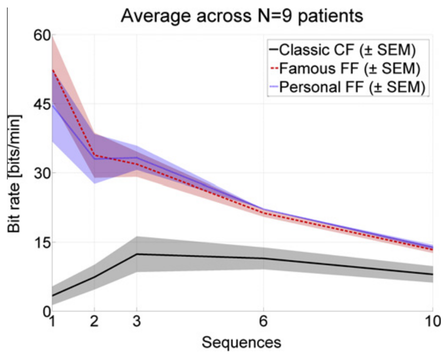


Fig. 4. Average bit rate and standard errors of the mean (SEM) of $N = 9$ patients with neurodegenerative disease.

formed equally well as compared to healthy users (famous FF: $Z = \text{inf}$, $p > .99$; personally known FF: $Z = -0.21$, $p = .83$; Bonferroni adjusted alpha level: $\alpha = .0167$; comparison of single-sequence performance).

Importantly, using CF two patients (patients 6 and 7) were not able to communicate with more than 40% online accuracy in any of the online sessions. We thus investigated if this high inefficiency was due to a bad calibration session (e.g. due to lack of attention in this critical period of the experiment) and could be improved by recalibrating the system. Therefore we recalibrated offline using data from the first online session (10 sequences, 5 characters). We then tested the recalibrated classifier weights on data from all other online sessions. Re-estimated performance still remained comparably low in the CF condition (OS_2–OS_5: [20%, 20%, 20%, 0%; patient 6], [40%, 0% 0%, 20%; patient 7]). We thus concluded that these patients are confronted with high BCI inefficiency when using CF, i.e. the BCI did not constitute a means of communication for these patients. In contrast, when using FF paradigms, patient 6 spelled with an average accuracy of 80% (famous FF) and 84% (personally known FF) whereas patient 7 did not make an error in any of the online sessions.

For patient 4, the difference between classification accuracy estimated offline and performance achieved online was particularly pronounced. With CF stimulation his offline performance was estimated 100% with 6 sequences whereas his best performance online was 60% with 10 sequences. The difference emerged largely due to his artificial ventilation that caused irregular contamination of the EEG signal. Thus, reducing the number of sequences entailed increasing the impact of artifacts. Consequently, patient 4 benefited strongly from the increased signal-to-noise ratio in the FF conditions, i.e. even with single-sequence classification he was as good as with 10 sequences using CF.

3.2. Comparison of event-related potentials

Fig. 5 displays the grand average ERPs of the patients for all three stimulus conditions from 0 to 500 ms post-stimulus. Grand average amplitudes and latencies for the N170, P300 and N400f are provided in Table 4. In line with previous results from healthy users (Kaufmann et al., 2011; and replicated in this study, see Fig. 6), face specific components were evoked in patients that increased the signal-to-noise ratio post-stimulus (N170, N400f; see Fig. 5). Difference in amplitude, however, did not reach significance (N170: famous FF vs. CF, $p = .19$; personally known FF vs. CF: $p = .03$; N400f: famous FF vs. CF, $p = .09$; personally known FF vs. CF: $p = .49$). Interestingly, P300 amplitudes were significantly increased in patients using FF paradigms ($p < .0028$ for both FF

898 T. Kaufmann et al. / Clinical Neurophysiology 124 (2013) 893–900

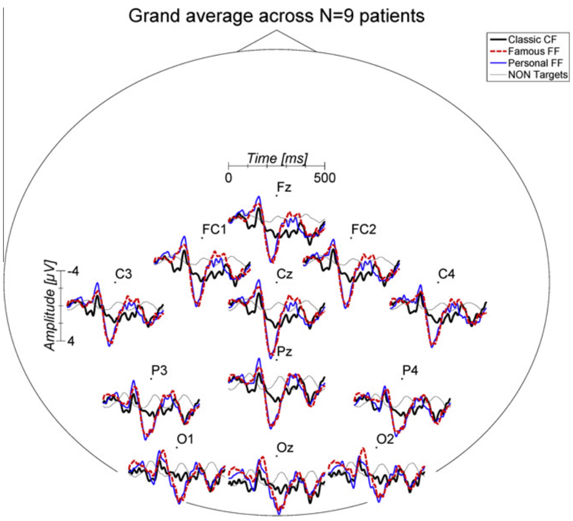


Fig. 5. Grand average topographic plot of all 12 electrodes for $N = 9$ patients with neurodegenerative disease. Data obtained from calibration sessions.

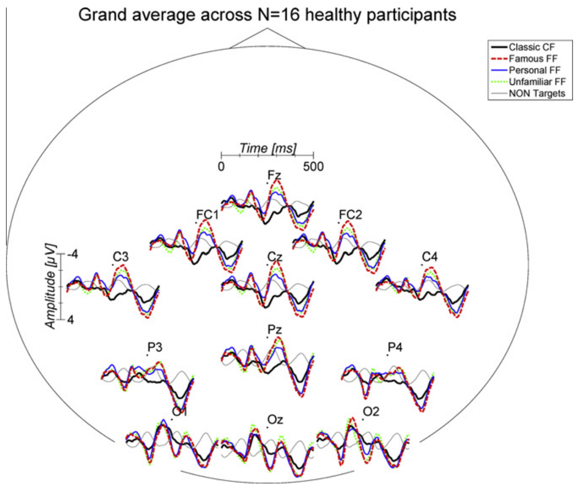


Fig. 6. Grand average topographic plot of all 12 electrodes for $N = 16$ healthy participants. Data obtained from calibration sessions.

conditions compared to CF; Bonferroni adjusted alpha level: $\alpha = .0167$).

We further assessed the impact of face familiarity on ERPs in the context of the BCI paradigm. Fig. 6 compares all four conditions in healthy participants (see also Table 4 for ERP amplitudes and latencies). The unknown face condition also evoked a face specific N400f of high amplitude. The difference in amplitude in the time interval of the N400f was significant between conditions ($H(3) = 15.1$, $p < .0017$). Post-hoc tests revealed no difference between the FF conditions, but all of them differed significantly from CF. Similarly, significant differences were found for the N170 ($H(3) = 13.7$, $p < .0033$) with no significant difference between FF conditions but significant differences between CF and all FF conditions. In healthy participants, P300 amplitudes were not significantly affected by the choice of stimulus condition ($H(3) = 4.04$, $p = .25$).

4. Discussion

Online ERP–BCI performance significantly increased when superimposing the characters of the matrix with faces. As known from several studies, patients performed significantly lower than healthy subjects in the classic character flashing condition. Not so, however, in the flashing faces condition.

4.1. Impact of stimulus material on BCI inefficiency

Our results confirm and extend previously reported benefits of the FF paradigm and show even larger effects in an online setting as compared to the results obtained from offline analysis (Kaufmann et al., 2011). Importantly, face stimuli effectively helped to overcome BCI inefficiency in patients with motor impairment due to neurodegenerative disease. All patients were able to spell with at least 60% accuracy in sessions with single-sequence classification using any of the FF paradigms. In contrast, when confronted with the classic CF approach, all of them were highly inefficient in single-sequence classification. Moreover, with CF three patients were not able to reach sufficient accuracy levels in any of the online sessions.

Our data provide evidence, that the increased signal-to-noise ratio in ERPs following FF stimuli may help to compensate for artifact contaminated EEG data, e.g. due to artificial respiration. But also BCI inefficiency due to other causes may be eliminated, e.g. as in the case of patient 7. He performed worst in the CF paradigm as compared to all other patients. In contrast, he also performed best of all in both FF paradigms. ERP analysis revealed that ERPs following face stimuli were very prominent in this patient whereas the classic CF paradigm elicited a small P300 only. Consequently BCI inefficiency was fully prevented in his case and

Table 4
Grand average ERP amplitudes and latencies.

	N170				P300				N400f			
	Amplitude (µV)		Latency (ms)		Amplitude (µV)		Latency (ms)		Amplitude (µV)		Latency (ms)	
	Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM
Patients (N = 9)												
Classic CF	−1.96	1.29	164.49	5.24	3.46	0.40	261.50	9.95	−0.82	0.69	361.11	14.16
Famous FF	−2.77	0.64	155.38	5.71	7.33	0.60	228.51	6.81	−2.70	0.73	340.49	11.90
Personal FF	−3.60	0.72	151.26	4.03	7.12	0.69	235.89	10.45	−1.63	0.70	335.94	13.06
Healthy (N = 16)												
Classic CF	−0.72	0.21	158.70	7.50	3.17	0.30	252.69	6.02	−0.43	0.29	339.60	9.02
Famous FF	−2.04	0.37	157.71	4.86	3.58	0.58	226.68	4.26	−4.10	0.80	306.60	5.46
Unfamiliar FF	−2.35	0.36	158.32	5.41	3.55	0.54	234.62	5.89	−2.35	0.59	300.29	4.41
Personal FF	−2.07	0.34	164.67	4.10	3.45	0.56	226.56	4.45	−3.02	0.66	307.13	4.66

Amplitudes and latencies were computed within pre-defined time windows (N170: 120–200 ms; P300: 200–300 ms; N400f: 280–400 ms; see Section 2.5) and are provided as grand average over participants. SEM, standard error of the mean; CF, character flashing; FF, face flashing.

even single-sequence classification was 100% accurate. These results strongly encourage the view that overcoming BCI inefficiency is a task for the BCI developer rather than for the BCI user.

Besides its large effect on BCI inefficiency, FF serves the goal of substantially increasing spelling bit rate. Recently, Zickler and colleagues (2011) reported "effectiveness", "reliability" and "speed" as being the major needs and requirements for BCIs according to the evaluation of potential BCI end-users. Our FF paradigm well addresses these issues.

4.2. Influence of face familiarity on performance

In our previous study we used famous faces for stimulation (Kaufmann et al., 2011) whereas Zhang and colleagues (2012) used unfamiliar faces or faces with loss of configural information. Jin and colleagues (2012) used one face of a fellow student known by all participants. Our study further investigated the effect of face familiarity on spelling accuracy. No significant effect was found between personally known and famous (Einstein) faces in patients. Even for unfamiliar faces (as tested in the healthy participant group) performance did not significantly differ from the familiar face stimulus conditions. In line with our expectations, all FF stimuli elicited an N170, i.e. a face-specific potential (Bentin et al., 1996) involved in rapid perception of faces (Eimer, 2000). Unexpectedly, an N400f was also found for unfamiliar faces. As the N400f is involved in face recognition and is usually observed for familiar faces (Eimer, 2000) we speculate that in the unfamiliar face condition it was elicited due to repetitive presentation of the same unfamiliar face. Therefore, primary familiarity of the face seems not to play an important role. In other words, a repetitively presented unknown face may rapidly become familiar, and thus, elicit the ERPs necessary to increase performance. In sum, we found no evidence that performance relies on primary face familiarity and thus, choice of appropriate face stimulus material may be according to user preference.

As can be seen from the grand average ERP in Fig. 6, N400f potentials were decreased for the personally known faces as compared to the famous face. Although this difference was not significant, it is not in line with the reported increase found by Touryan et al. (2011). We speculate that the quality of the stimulus material may explain such decreased amplitude. Individually tailored stimulus material entailed a compromise in stimulus quality. Several photographs were of low resolution, bad exposure or artificial color. In parts we were able to correct for these issues using image editing software. However, compared to the high quality of the Einstein face stimulus (famous FF), recognition of personally known face stimuli might – at least in some cases – have been difficult in the framework of rapid stimulus presentation.

4.3. Limitations

Generalization of our results is limited in that all participants spelled one word per session only, i.e. one time the word "BRAIN" per number of sequences. We chose this study design to limit the burden for the patients. Our design comprised five times five letters for each participant and condition. The influence of single errors on overall performance was thus relatively high – but apparently equally high for all conditions. Importantly, the strong benefit of all FF paradigms was nevertheless clearly visible, as overall error rate with CF was higher in eight of nine patients and in 13 of 16 healthy participants; the remainder of subjects achieved equal or better performance with at least one of the FF paradigms.

Further research is needed to assess the impact of FF stimuli on BCI performance in a larger group of patients. In this study we chose a heterogeneous sample including a wide age range and different types of disabilities with different degrees of impairment. Our results confirm great benefit of the FF paradigm in eight of nine patients and do not indicate that age or the degree of physical impairment influence performance. However, to systematically investigate these factors, a larger sample is needed, preferably including locked-in patients. Also, patients with diseases at different levels of cortical involvement would be needed to investigate if enhanced spelling accuracy in the FF condition depends on the ability to recruit cortical resources. In this regard it appears interesting to consider that processing of faces is in large parts considered an automatic (i.e. resource independent) task.

Finally, although our patient sample comprised severely motor impaired subjects, all diagnosed with neurodegenerative disease, no locked-in patient was included. However, it is well known that when ALS is diagnosed, already around 30% of the motoneurons and also some cortical neurons are degenerated (Swash and Ingram, 1988). It is highly encouraging that these patients in our sample benefited even more than those with spinal muscular atrophy where upper motoneurons and other cortical neurons are not affected. The FF paradigm is thus a promising perspective for increasing bit-rate in locked-in patients with intact vision.

5. Conclusion

This study compared classic target flashing against paradigms in which targets were transparently overlaid with faces. All face stimuli resulted in significantly increased online accuracy compared to classic character flashing. Patients benefited from the flashing face conditions to such an extent that their performance in single-sequence online classification did not significantly differ from that achieved by healthy users. This allowed for substantially decreasing the number of character flashes, and thus time needed for spelling while maintaining high performance levels. Importantly, the increased signal to noise ratio in ERPs following face stimuli effectively prevented BCI inefficiency in patients with neurodegenerative disease. Thus, we strongly encourage exploiting the advantage related to the use of faces in ERP–BCI controlled applications, in particular when working with severely motor-impaired end-users.

Acknowledgments

The authors declare no competing financial interests. This work was supported by the Deutsche Forschungsgesellschaft (DFG) RTG 1253/1 and the European ICT Programme Project FP7-224631. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

References

Bentin S, Allison T, Puce A, Perez E, McCarthy G. Electrophysiological studies of face perception in humans. J Cogn Neurosci 1996;8:551–65.

Blankertz B, Sannelli C, Halder S, Hammer EM, Kübler A, Müller K-R, et al. Neurophysiological predictor of SMR-based BCI performance. Neuroimage 2010;51:1303–9.

Eimer M. Event-related brain potentials distinguish processing stages involved in face perception and recognition. Clin Neurophysiol 2000;111:694–705.

Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalogr Clin Neurophysiol 1988;70:510–23.

Jin J, Allison BZ, Kaufmann T, Kübler A, Zhang Y, Wang X, et al. The changing face of P300 BCIs: a comparison of stimulus changes in a P300 BCI involving faces, emotion, and movement. PLoS One 2012;7(11):e49688.

Kaufmann T, Schulz SM, Grünzinger C, Kübler A. Flashing characters with famous faces improves ERP-based brain–computer interface performance. J Neural Eng 2011;8:056016.

Kleih SC, Kaufmann T, Zickler C, Halder S, Leotta F, Cincotti F, et al. Out of the frying pan into the fire – the P300-based BCI faces real-world challenges. Prog Brain Res 2011;194:27–46.

Krusienski DJ, Sellers EW, McFarland DJ, Vaughan TM, Wolpaw JR. Toward enhanced P300 speller performance. J Neurosci Methods 2008;167:15–21.

Kübler A, Neumann N, Kaiser J, Kotchoubey B, Hinterberger T, Birbaumer NP. Brain–computer communication: self-regulation of slow cortical potentials for verbal communication. Arch Phys Med Rehabil 2001;82:1533–9.

Kübler A, Müller K-R. Introduction to brain–computer interfacing. In: Dornhege G, Millan JDR, Hinterberger T, McFarland D, Müller K-R, editors. Towards brain–computer interfacing. Cambridge, MA: MIT Press; 2007.

Kübler A, Birbaumer N. Brain–computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralyzed patients? Clin Neurophysiol 2008;119:2658–66.

Kübler A, Blankertz B, Müller K-R, Neuper C. A model of BCI-control. In: Proceedings of the fifth international BCI conference, Graz, Austria, 2011.

Mak JN, Arbel Y, Minett JW, McCane LM, Yuksel B, Ryan D, et al. Optimizing the P300-based brain–computer interface: current status, limitations and future directions. J Neural Eng 2011;8:025003.

Nijboer F, Sellers EW, Mellinger J, Jordan MA, Matuz T, Furdea A, et al. A P300-based brain–computer interface for people with amyotrophic lateral sclerosis. Clin Neurophysiol 2008;119:1909–16.

Pierce JR. An introduction to information theory. New York, NY: Dover Press; 1980.

Polich J. Updating P300: an integrative theory of P3a and P3b. Clin Neurophysiol 2007;118:2128–48.

Schalk G, McFarland DJ, Hinterberger T, Birbaumer N, Wolpaw JR. BCI2000: a general-purpose brain–computer interface (BCI) system. IEEE Trans Biomed Eng 2004;51:1034–43.

Sellers EW, Donchin E. A P300-based brain–computer interface: initial tests by ALS patients. Clin Neurophysiol 2006;117:538–48.

Shannon CE, Weaver W. The mathematical theory of communication. Urbana, IL: University of Illinois Press; 1964.

Swash M, Ingram D. Preclinical and subclinical events in motor neuron disease. J Neurol Neurosurg Psychiatry 1988;51:165–8.

Touryan J, Gibson L, Horne JH, Weber P. Real-time measurement of face recognition in rapid serial visual presentation. Front Psychol 2011;2:42.

Vidaurre C, Blankertz B. Towards a cure for BCI illiteracy. Brain Topogr 2010;23:194–8.

Wolpaw JR, Ramoser H, McFarland DJ, Pfurtscheller G. EEG-based communication: improved accuracy by response verification. IEEE Trans Rehabil Eng 1998;6:326–33.

Zickler C, Riccio A, Leotta F, Hillian-Tress S, Halder S, Holz E, et al. A brain–computer interface as input channel for a standard assistive technology software. Clin EEG Neurosci 2011;42:236–44.

Zhang Y, Zhao Q, Jin J, Wang X, Cichocki A. A novel BCI based on ERP components sensitive to configural processing of human faces. J Neural Eng 2012;9:026018.

3 General Discussion

In chapter 1.5 basic issues were identified that built the basis of this dissertation project:

1. Can we identify performance predictors so that we can provide users with individual BCI solutions without the need of multiple, demanding testing sessions?
2. Can we provide complex BCI technology in an automated, user-friendly and easy-to-use manner, so that BCIs can be used without expert support at end-users' homes?
3. How can we account for and improve the low information transfer rates as compared to other existing assistive technology solutions?
4. How can we prevent the performance drop often seen when bringing BCI technology that was tested in healthy users to those with severe motor impairment?

As described in chapter 1.5 and depicted in figure 9 several aspects in the model of BCI control (Kübler et al., 2011) were thus under investigation to address all of these issues in several studies.

In the following it is discussed how and to what extend these issues could be solved. Results are evaluated in the context of recent literature. Furthermore, limitations as well as future experimentations are discussed.

1. **A peripheral-physiological performance predictor for ERP-BCIs was identified.**

   This thesis identified RSAnorm, an index of cardiac vagal activation and thus of inhibitory control capacity, as a peripheral-physiological predictor of ERP-BCI performance, accounting for almost 26% of variance (Kaufmann et al., 2012a).

   RSAnorm is computed by normalizing high frequency power estimates of HRV. Higher values of RSAnorm reflect higher vagal activation. It was found to be less prone to sympathetic influences than a simple high frequency power estimate (factor 6.6; Hedman et al., 1995; for a review, e.g., Grossman and Taylor, 2007). As heart rate variability is usually indexed by a plethora of parameters in the time and frequency domain, it is important to note that not only RSAnorm but also other HRV parameters significantly correlated with performance, thereby displaying the same relationship. Correlation of time domain parameters was lower as expected, however this might be due to remaining sympathetic influences in these parameters. Yet that correlation of RSAnorm was stronger than of non-normalized high frequency estimates supports the hypothesized relationship between cardiac vagal activation and ERP-BCI performance. The predictive value was further strengthened by computing a robust regression to account for outliers in the data. As the robust regression model did not differ from a simple least square regression model, it can be concluded that the strong correlation was not due to outliers.

Patients were not yet targeted in this basic investigation, as identification of predictor variables requires large ($\geq$30) sample sizes. Consequently, predictors are more easily assessable in healthy samples, thereby also limiting patient effort. However, future research needs to investigate if the identified relation between RSAnorm and BCI performance is present in patient samples as well.

HRV is particularly easy to assess. It is computed from a 5-min resting period recording of heartbeats. Either an ECG can be obtained from a small number of electrodes (e.g. 3 electrodes in lead II configuration) or inter-beat intervals can be directly recorded using a heart rate monitor such as a sport watch (e.g. Polar S810; Nunan et al., 2009). Consequently, when predicting BCI performance based on HRV, no EEG assessment is necessary. This may be of high practical value, especially for patients for whom washing hair is of high effort (Zickler et al., 2011). Different BCI paradigms usually require different electrode setups. Thus, testing different BCI systems usually requires different testing sessions. Starting with the system that is predicted to be most likely controllable without the need of obtaining EEG signals could reduce this effort. From the results of this thesis, it is concluded that HRV could fulfill this criteria. On the other hand, the high effort in EEG preparation may not be an exclusion criterion if stronger predictors than HRV are identified. Recent literature suggests ERP-BCI prediction based on EEG data obtained from an auditory oddball paradigm and reports correlations up to r=.68 (Halder et al., 2013; for comparison, correlation between RSAnorm and BCI performance was r=.506). The practical value of HRV as a performance predictor thus has to be further investigated. As done by Halder and colleagues (2013) other modalities could be assessed as well, e.g. auditory and tactile ERP-BCIs. Such BCIs may require more inhibitory control capacity, as all stimuli (targets and non-targets) are inevitably perceived. In the visual ERP-BCI it is possible to focus on the target stimulus and ignore non-target stimuli by simply avoiding looking at them. Such suppression of irrelevant stimuli is not as easily possible in other modalities. HRV could thus potentially display a stronger predictive value in these BCI modalities.

Apart from their value as a predictor variable, correlates of BCI performance could be useful in describing participant samples that have been integrated in studies. For example, if a study reports an achieved improvement based on results from a small sample size, generalization of these results to a larger population is often questionable. Yet, if one would describe the outcome of their personal predictor estimates, the reader could potentially estimate what BCI accuracy is to be expected from such participants. A participant performing with high accuracy, although the estimated performance based on his/her predictor variable is low, would underline a suggested improvement more than a participant with an already expected high performance level. HRV estimates of vagal activation have been described as reliable indices if obtained from at least two recordings (to reduce occasion specific, cardiac effects; Bertsch et al., 2012). If, for example, conducted in two recordings prior to and after a BCI session, HRV could thus potentially serve for describing the expected potential of participants.

An analysis software called *ARTiiFACT* for processing of artifacts in ECG and IBI data was developed (Kaufmann et al., 2011c). It particularly targets recordings where either no ECG is available (e.g. recording from heart rate monitors) or where ECG data is contaminated so that no peak detection is possible. The fact that even minor artifact contamination of data has a high impact on the computed HRV renders artifact processing of utmost importance (Berntson and Stowell, 1998; figure 10). The effect of a single artifact on computed HRV may be greater than any true physiological variance in the data (Berntson and Stowell, 1998). Yet publications on HRV often conceal their artifact processing methods, therefore giving reason to believe that in fact no artifact processing has been done. A possible reason might be the lack of software implementations of artifact processing algorithms.

Berntson and colleagues (1990) theoretically proposed an algorithm for artifact detection that was implemented in the framework of this thesis using MATLAB (The Mathworks Inc., USA). Validation against the artifact processing methods implemented in the commonly used HRV analysis software Kubios (Niskanen et al., 2004; Tarvainen et al., 2009) proved the superiority of this algorithm (Kaufmann et al., 2011c). To provide the method to the research community, the algorithm was incorporated into an easy-to-use graphical user interface. Apart from artifact detection in IBI data, *ARTiiFACT* allows for processing of ECG data, processing of the detected artifacts in IBI data, computing HRV and statistically validating data. Thus, it targets the full analysis cascade from raw data to computation of HRV parameters.

ARTiiFACT has been highly accepted by users. To date a copy of the software has been requested by more than 100 researchers from all over the world and freely shared through dropbox (www.dropbox.com). Development is ongoing to incorporate new features and to handle user requirements and suggestions. An update to version 2.0 was released in April 2013. It incorporates new features such as batch processing options to handle a large amount of data, an improved graphical user interface (figure 11) and new processing options for ECG and IBI data. Due to the high demand, a website was developed that was released with the new version of *ARTiiFACT* (www.artiifact.de). It incorporates training video tutorials for software handling and explanation of the analysis procedures. Frequently asked questions are addressed on the website, yet ongoing support will be provided via email. After a quick registration procedure, a download repository is freely accessible. It allows for access of different versions of *ARTiiFACT*, e.g. for different software platforms or different release versions.

A

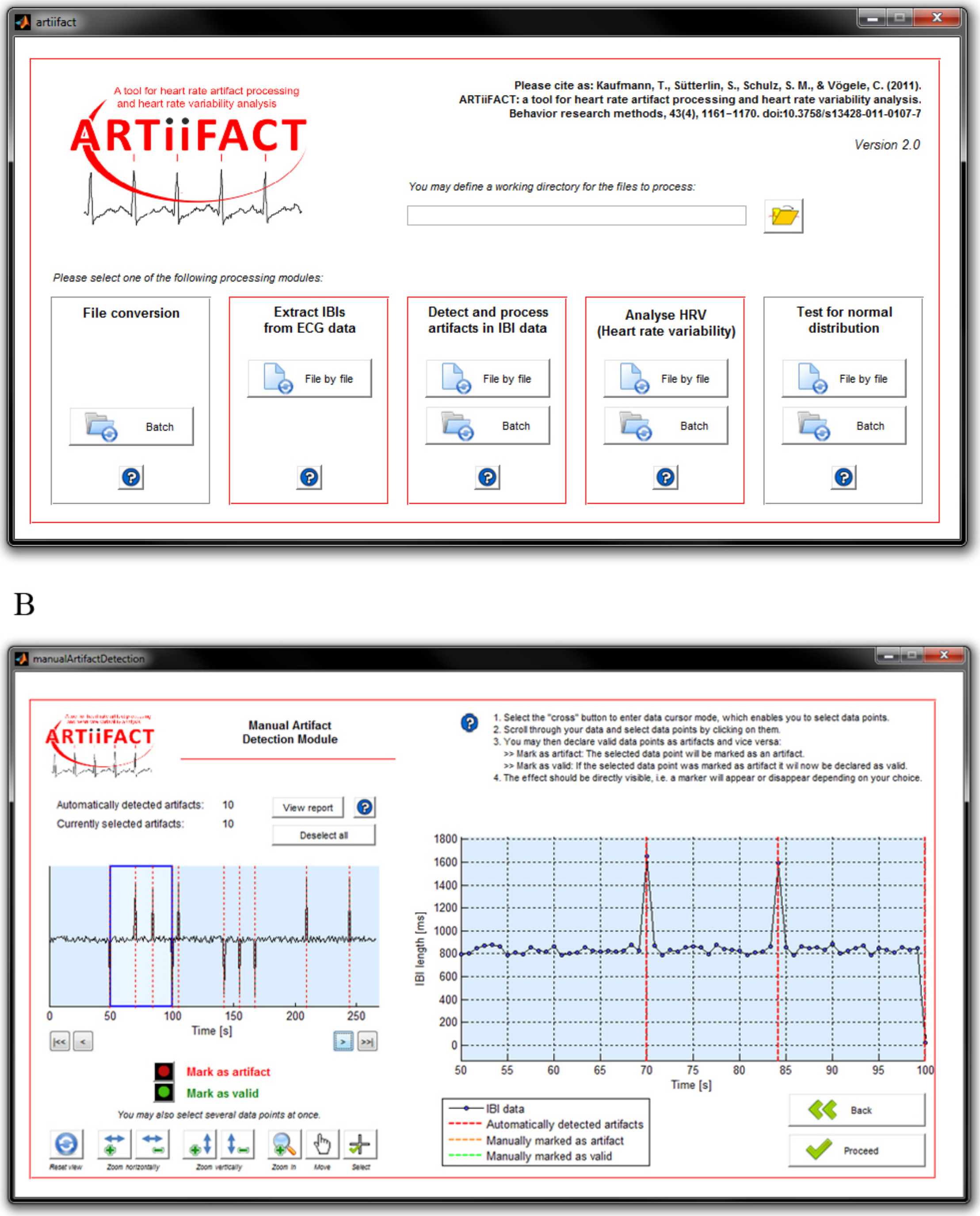


Figure 11: New release of *ARTiiFACT* (version 2). Apart from an improved graphical user interface it incorporates new features such as options for batch processing. The figure provides two exemplary screenshots of (A) the main module and (B) the module for artifact detection. More details are available on www.artiifact.de.

2. **Complex BCI technology was integrated into a user-friendly and easy-to-use interface. The calibration procedure and parameterization of the software was automatized so that requirement for expert support was reduced.**

   Decreasing user effort and dependence on expert support touches aspects on both the hard- and software side. In this thesis the software side was addressed (Kaufmann et al., 2012b). In particular, this study targeted handling of the BCI software and automatizing necessary analysis procedures thereby reducing the requirement for expert support during BCI use. All naïve participants were able to handle the software (referred to as *Optimized Communication Interface*) on their own after a quick tutorial. This result renders control by caregivers and/or relatives possible. This is of particular importance in the changing clinical environment setting that patients are often confronted with, as no intensive training in EEG analysis and handling of communication devices can be expected from patients' caregivers and/or family members. Their time is often limited to nursing and organizational tasks. A system that substitutes an expert by analyzing data automatically, judging its quality and deciding on which steps to be carried out next may be of high value in this setting.

   The study proofed the concept of expert independent BCI control feasible, i.e. software handling (set-up) can be reduced to a minimum number of straightforward operations and data analysis can be automatized. Yet regarding actual use of the device for communication in daily life situations, it faces the limitation that all BCI users were healthy participants. In patient samples one often has to deal with external artifacts (e.g. ventilation device or other clinical devices), internal artifacts (e.g. due to spasm; figure 12 for an example) and with psychological issues (e.g. attention allocation or workload difficulties). This thesis provided a means to account for lower signal-to-noise ratios in the recorded EEG (Kaufmann et al., 2011b, 2013b). The therein-suggested new stimulus material was recently integrated into the *Optimized Communication Interface* and is currently used in patient testing.

   However, various issues may occur that cannot automatically be dealt with and thus support during BCI use may still be required. Recent literature suggested a BCI in home-use in which support by experts was delivered via remote internet access (Holz et al., 2013).

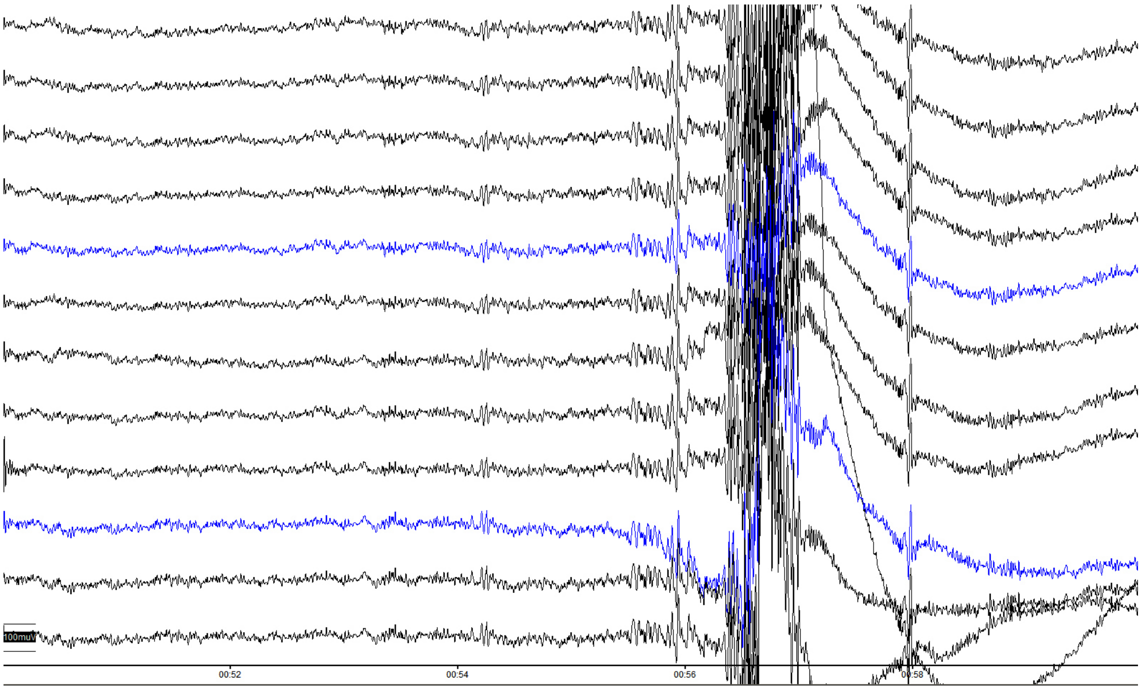


Figure 12: Exemplary illustration of an artifact contaminated EEG due to spasm. The contamination affects signals at all electrode sites. It induces baseline drifts so that several seconds of EEG recording are fully unreliable. Data was obtained from a patient with cerebral palsy visited in the period of this dissertation project.

The authors installed a BCI system at the home of a 72-year old locked-in patient and trained the family and caregivers to set-up an EEG and handle the software. The system was calibrated once, i.e. the generated classifier was used in all further BCI sessions. Ongoing support was then provided through remote access. After each of the 86 BCI sessions that the patient conducted with support of her family and caregivers, data was automatically streamed to a server. As such, the authors were able to overview the ongoing activity in BCI use, to analyze data in case of problems or even to modify the BCI system by means of remote computer access. Their evaluation results show that it is possible to provide sufficient expert support remotely. Sources of dissatisfaction were technical problems and low accuracy, the latter potentially due to misplacement of the EEG cap or insufficient preparation of electrodes.

One possibility to account for misplacement of the EEG cap is an auto-calibration as proposed in this thesis (Kaufmann et al., 2012b). After EEG preparation, a short calibration session would individually adjust classifier weights to the current EEG set-up. To prevent unnecessary time effort, the auto-calibration could also be made available as an optional mode, so that caregivers (or even automatically the BCI itself) could start the recalibration if insufficient control accuracy was achieved in the beginning of a BCI session. In addition, recent literature suggests recalibrating the system if the backspace key of the communication application (as an indicator for low performance) is used too frequently (Daucé and Proix, 2013).

This thesis addressed complexity of the software but not of the hardware setup. Currently, reliable EEG signals are obtained from wet electrodes. The preparation effort and the required washing hair after BCI use was rated highly uncomfortable by potential BCI users (Zickler et al., 2011). Research as well as industry are currently developing new systems that do not require electrode gel, i.e. systems using water-soaked electrodes (e.g. Emotiv Epoc) or dry electrode systems (e.g. gTec gSahara; Starlab Enobio). For a review on different recording techniques, see e.g. Mak and colleagues (2011). An extensive testing of the EU project TOBI revealed issues with signal-to-noise ratios of dry compared to wet electrode systems (D8.12, 2013; TOBI Project; www.tobi-project.org). For example the g.Sahara system was prone to external artifacts despite the fact that the quality of the recorded signal was good. The Enobio system on the other hand was less prone to artifacts, yet the low signal-to-noise ratio may restrict its practical use to very strong EEG signal.

Whereas signal quality is still an issue to be addressed, high cost of EEG systems has already decreased to an affordable amount. Debener and colleagues (2012) modified a commercially available, low cost EEG system from Emotiv (www.emotiv.com) in that they replaced the electrodes with high quality wet electrodes, but used the built-in amplifier of the Emotiv system. As such, they obtained an EEG system of high quality but low cost. From a healthcare perspective, this is clearly promising as BCIs may become financially competitive to other assistive technology.

3. **Information transfer rates were significantly improved by modification of the stimulus material used in ERP-BCIs. Furthermore, information throughput was enhanced with integration of predictive text suggestions into the matrix.**

   This thesis introduced face stimuli to ERP-based BCIs (Kaufmann et al., 2011b). This new face paradigm additionally elicits specific ERPs involved in recognition and processing of faces thereby increasing the signal-to-noise ratio of the EEG post-stimulus. In a proof-of-concept study the paradigm was first evaluated offline with healthy participants (Kaufmann et al., 2011b). From the literature it was assumed that the results might transfer well to patients as face recognition in ALS patients was reported to remain intact (Abrahams et al., 1996). Yet, to allow for a thorough investigation of the paradigm's potential benefit in patient samples, a study was conducted with N=9 patients and N=16 healthy controls (Kaufmann et al., 2013b).

Both studies clearly demonstrated superiority of the proposed face paradigm over the classic character-highlighting paradigm that has been used in almost all visual ERP-BCI studies since its proposal 25 years ago (Farwell and Donchin, 1988). Most research tried to enhance information transfer rates by improving signal processing techniques or modifying procedural aspects of the oddball paradigm, such as stimulus pattern or stimulus timing (for review, e.g., Kleih et al., 2011; Mak et al., 2011). Martens and colleagues (2009) changed the stimulus material introducing bars that flip horizontally for eliciting ERPs. The authors demonstrated that this modification reduced refractory effects of ERPs and led to a significant performance gain in two of six participants. Consequently, target-to-target intervals could be reduced without negatively inducing refractory effects. However, such timing can be decreased only to some extent, thereby limiting the practical benefit of this approach. In addition, fast stimulation may annoy users as shown for SSVEP BCIs (Allison et al., 2010).

Face stimuli as introduced in this thesis have several advantages:

- Above all, they elicit additional, distinct ERPs thereby increasing the signal-to-noise ratio of the EEG post-stimulus.
- They are large, bright stimuli that can easily be recognized.
- They are ”odder” than simply highlighting a target or changing its color, as they differ from the character that users focus on.

The latter two aspects may positively affect attention allocation capabilities and may ease recognition of stimuli in case vision is impaired. The increased signal-to-noise ratio resulted in significantly higher classification accuracies.

Consequently online information transfer rates achieved with the face paradigm were high as compared to the classic character-flashing paradigm (Kaufmann et al., 2013b). Importantly, even when reducing the number of stimulation sequences to the minimum of one sequence, several users were able to spell without any error, i.e. these users spelled highly reliably at maximum speed in an online setting. Thus, stimulus material displayed a strong effect on information transfer rates.

Further research is required to identify the best stimulus properties of the face stimuli. Recently, Jin and colleagues (2012) investigated the effect of face emotion and face motion on performance, however, no difference to the neutral face condition was found. This thesis found no effect of face familiarity on BCI performance, i.e. unknown, famous or even personally known faces resulted in similar classification accuracies (Kaufmann et al., 2013b). From the literature on face processing it was expected that personally familiar faces elicit a higher N400f than famous faces (Touryan et al., 2011) and that both conditions display a higher N400f than the unknown face condition. However, the results in this thesis were (to some extend) opposite. It is possible that the quality of the stimulus material for personally known faces was too low therefore constraining fast recognition during the stimulation process. The unknown face on the other hand may rapidly have gained familiarity due to repetitive presentation. However, the effect of face familiarity appears to be marginal, as performance did merely differ between conditions. A more promising modification of the face paradigm appears to be an inversion of the presented face (Zhang et al., 2012). The authors reported increased information transfer rates following such inverted stimuli.

Apart from targeting the signal quality of the elicited ERPs, this thesis increased information transfer rates by integrating predictive text entry (PTE) into the BCI system (Kaufmann et al., 2012b). A prior work by Ryan and colleagues (2011) reported decreased performance when spelling with predictive text suggestions. In their implementation, users had to screen a list of predictive text suggestions, keep in mind a number associated with the intended text suggestion and focus on this corresponding number in a separate stimulation matrix. The reported decrease in performance was potentially due to this indirect approach. When integrating text suggestions directly into the matrix, thereby allowing selecting them in exactly the same manner as selecting characters, no performance drop was visible (Kaufmann et al., 2012b). Raw information transfer rates (amount of correctly delivered commands per time) were comparably high to spelling without PTE. True information transfer rates (amount of correctly communicated characters per time) were higher with PTE for all participants. The latter depends on the length of the spelled words and thus has to be interpreted with caution, yet it displays the potential of PTE in BCI applications. In this implementation, the PTE system only suggested text based on a prior ranking of the word data base, independent of contextual sense or grammar rules. It could thus be further improved when integrating grammar rules or context based filters.

Another option for increasing information transfer is to adjust the number of stimulation sequences dynamically based on the certainty of the classification output. This procedure, referred to as dynamic stopping, is thus capable of accouning for trial-to-trial differences. For example, if after few stimulations a certain decision can already be made, further stimulation would unnecessarily increase spelling time. On the other hand, if after a certain number of stimulations no decision can be made further stimulation may prevent false decision. Different algorithms have been proposed to implement such stopping method (e.g., Zhang et al., 2008; Lenhardt et al., 2008; Höhne et al., 2010; Liu et al., 2010; Jin et al., 2011; Schreuder et al., 2011b) and are usually superior to the fixed sequence approach (Schreuder et al., 2011a).

Importantly, as none of the discussed approaches is exclusive to another, these approaches could be merged in a practical BCI application. Such BCI would incorporate face stimulation, base classification on a dynamically adjusted number of sequences and include text suggestions directly into the matrix.

4. **Performance drops or even BCI inefficiency was prevented in patients with neurodegenerative disease when confronted with the improved stimulus material.**

   Face stimuli did not only positively affect information transfer rates in online BCI use. This thesis also displayed its impact on BCI inefficiency (Kaufmann et al., 2013b). When using the classic character-flashing paradigm, three patients did not reach sufficient accuracy levels. Not even with a high number of stimulation sequences did two of them achieve accuracies above 40%. By contrast, all of them performed well using the face paradigms. Using face stimuli, one did not even make a single error in any of the spelling sessions. As the bad results achieved with the classic paradigm could be due to bad calibration runs, classifiers were recomputed from online data. Yet, performance remained comparably low. Consequently it is concluded that these users were inefficient with regard to the commonly used character-highlighting paradigm. Face stimuli, on the other hand, effectively helped to overcome this inefficiency phenomenon.

This finding is further strengthened when comparing performance of patients to that of healthy controls. A significant performance drop was visible for patients when exposed to the classic character-flashing paradigm. This is in line with previous reports (e.g., Piccione et al., 2006). However, when comparing performance achieved with the face paradigms, patient samples did not significantly differ from healthy controls.

One limitation of this study is that all patients had retained muscle control and all of them had a reliable means of communication other than a BCI. However, disease progression was advanced for several of the included patients. Recent research investigated the effect of ALS disease severity and disease progression on BCI performance (Silvoni et al., 2013). The authors reported that learning of BCI control is possible in all stages of the disease (except for CLIS where no reliable communication has ever been established; Kübler and Birbaumer, 2008; Birbaumer et al., 2012). As eye-gaze control was found to be a critical factor for matrix based ERP-BCI communication (Treder and Blankertz, 2010; Brunner et al., 2010), use of the proposed face paradigm in patients with strongly limited gaze control may require modifications to reduce gaze dependence (so called gaze-"independent" speller paradigms; for review, Riccio et al., 2012). Still, such spellers can incorporate face stimuli (Kaufmann et al., 2013a) and may well display the same benefits as for the gaze-dependent matrix paradigm.

Apart from the actual validation results, this study further underlines the importance of testing proposed systems in patient samples. The strength of benefit may well differ between healthy users and patients. This is also the case for the recorded signals, internal and external artifacts (figure 12), the environmental settings that a BCI system has to deal with (data from healthy users are usually acquired in artificial laboratory environments) as well as the needs and requirements of users. For example, studies on gaze dependence of visual ERP-BCIs and studies proposing the above described gaze-independent spellers should target users with severe gaze-impairment. Although conditions of covert attention can be imitated by healthy controls, they do not replace valid testing with severely impaired users. Some patients for example suffer uncontrollable eye-drifts whereas others only have problems in fixating.

Another example of differences between healthy users and patients is the choice of modality in ERP-BCI use. Aloise and colleagues (2007) compared classification accuracies on ERPs elicited in three modalities, i.e. visual, tactile and auditory. Their results clearly displayed superiority of the visual modality. Yet in a case study with a patient in the classical LIS (after a brainstem stroke), tactile modality revealed to be much more reliable than visual and auditory modalities (Kaufmann et al., 2013a). Figure 13a displays ERPs elicited in these three modalities, indicating that tactile ERPs were most prominent. Figure 13b displays reliability of offline classification accuracy obtained by calibrating the system on one session and testing on another session (and vice versa). The tactile modality was far more reliable across sessions than other modalities. This case study thus demonstrates that research should not be narrowed to one direction (e.g. the visual modality) but retain different options to choose from (Kaufmann et al., subm).

Consequently, not only regarding scientific investigations, incorporation of patient samples is necessary. To allow for bringing BCI technology out of the lab to end-users' homes means finding the best solution for every individual user. A user-centered design approach involving an individual user from the early developmental process into implementation and testing of a BCI system is considered highly valuable (Zickler et al., 2011; Holz et al., 2012; Kübler et al., 2013; Holz et al., subm). Only their feedback can help to provide practical BCI systems. The future of BCI-based communication research will particularly depend on the proof of clinical usefulness of these devices (Wolpaw and Wolpaw, 2012b).

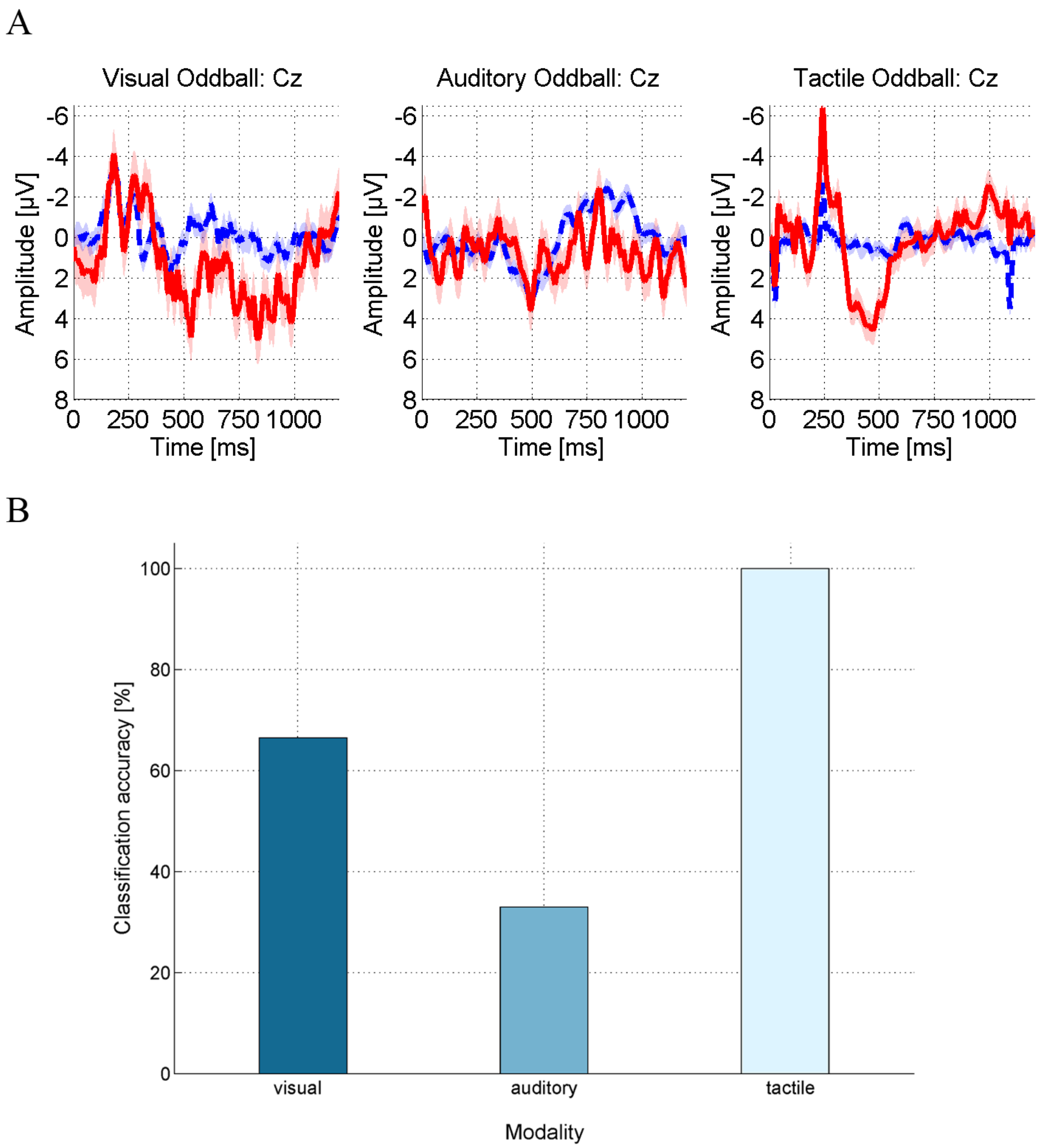


Figure 13: (A) ERPs in different modalities obtained in a case study with a patient in classical LIS (Kaufmann et al., 2013a). The tactile modality displayed most prominent and reliable ERP response. (B) Offline classification accuracy for different modalities. Two runs were used to obtain classification results. A classifier was trained based on one run and tested on the other - and vice versa. The figure depicts the average classification accuracy. Tactile modality displayed most reliable classification accuracy, i.e. both runs could be classified with 100% accuracy.

4 Conclusion and Outlook

Milestones were achieved in improving information transfer rates (Kaufmann et al., 2011b, 2012b, 2013b), preventing performance drops in patients with neurodegenerative disease (Kaufmann et al., 2013b) and providing BCI technology in an automatized and easy-to-use manner (Kaufmann et al., 2012b). A peripheral-physiological performance correlate was identified (Kaufmann et al., 2012a) encouraging the development of an artifact processing software for heart beat data to allow for its straightforward assessment (*ARTiiFACT*; Kaufmann et al., 2011c). The software is provided to the research community at no charge upon request. Taken together, this thesis addressed several aspects in the model of BCI control (Kübler et al., 2011) to improve current BCI-based communication technology and to enhance our understanding of physiological correlates of BCI performance.

The strong boost of BCI performance due to the use of face stimuli in ERP-BCI paradigms highly encourages its integration in practical BCI systems. Apart from its introduction and positive validation in healthy participants and patients in this thesis (Kaufmann et al., 2011b, 2013b), the strong benefit was confirmed by other studies (Zhang et al., 2012; Jin et al., 2012). Recently, face stimuli were integrated into the remote-support BCI system described earlier (Holz et al., 2013). The benefit is apparent, as the patient is now able to deliver commands at half the time required in earlier sessions (number of sequences reduced from 10 to 5). In conclusion, modification of stimulus material is a promising direction to be explored in future research and its use is of direct practical value for BCIs as a means of communication.

This thesis provided proof that complex BCI technology can be automatized and integrated into an easy-to-use interface to be handled by non-BCI-experts (Kaufmann et al., 2012b). The study demonstrates the potential of auto-calibrating systems such that all naïve participants were able to use the device for communicating sentences without expert support. However, it faces the limitation that all participants were young, healthy individuals and that the experiment was conducted in a laboratory environment. To prove its value in patients' daily life situations, a BCI system is required to cope with various issues that were not present in the framework of this proof-of-concept study. Future research is thus required to investigate methods for automatically dealing with internal and external artifacts, non-stationarity within and between sessions, insufficient EEG electrode placement as well as sudden hardware deficits (e.g. electrode breakdown).

Heart rate variability estimates of cardiac vagal activation as an index of inhibitory control were significantly related to performance in ERP-BCI use (Kaufmann et al., 2012a). HRV as a reliable index of inhibitory control, easy to assess from short heartbeat recordings (ECG or IBI), may thus allow for prediction of BCI performance and for explaining between-subject differences. Future research is required to replicate the findings in patients and to integrate different other physiological and psychological predictors into a prediction model for various BCI paradigms. The freely available artifact processing software *ARTiiFACT* (Kaufmann et al., 2011c) allows for easy yet reliable assessment of the predictor variable. This tool is also valuable for other psychophysiological research and has already been requested by many researchers worldwide. Development will be ongoing. Recently, a new version along with a website were released (www.artiifact.de).

In sum, this thesis successfully targeted BCI-based communication technology, thereby further paving the way for practical, clinically useful BCIs. BCI research is a growing field with exponentially increasing publication outcome since the early nineties (Wolpaw and Wolpaw, 2012b; Kübler et al., 2013). A strong potential of the field lies in the interdisciplinarity of its researchers' backgrounds. It is of utmost importance to emphasize the potential end-users, integrating them into the early stages of BCI research and development (Zickler et al., 2011; Holz et al., 2012; Kübler et al., 2013; Kaufmann et al., 2013a). Recent improvements of efficiency and reliability of BCI-based communication systems are promising, such that BCIs will in the future provide a clinically useful means of communication. Cases like our client described at the beginning of this thesis motivate spurring the steadily further development of brain-computer interface technology into practical systems for independent home-use.

References

Abrahams, S., Goldstein, L. H., Kew, J. J. M., Brooks, D. J., Lloyd, C. M., Frith, C. D., and Leigh, P. N. (1996). Frontal lobe dysfunction in amyotrophic lateral sclerosis a PET study. *Brain*, 119(6):2105–2120.

Abrahams, S., Leigh, P., Kew, J., Goldstein, L., Lloyd, C., and Brooks, D. (1995). A positron emission tomography study of frontal lobe function (verbal fluency) in amyotrophic lateral sclerosis. *Journal of the Neurological Sciences*, 129, Supplement:44–46.

Abresch, R. T., Seyden, N. K., and Wineinger, M. A. (1998). Quality of life. issues for persons with neuromuscular diseases. *Physical Medicine and Rehabilitation Clinics of North America*, 9(1):233–248. PMID: 9894142.

Allison, B., Luth, T., Valbuena, D., Teymourian, A., Volosyak, I., and Graser, A. (2010). BCI demographics: How many (and what kinds of) people can use an SSVEP BCI? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(2):107–116.

Allison, B. Z., Dunne, S., Leeb, R., Millán, J. D. R., and Nijholt, A. (2012a). *Towards Practical Brain-Computer Interfaces: Bridging the Gap from Research to Real-World Applications*. Springer, Biological and Medical Physics, Biomedical Engineering.

Allison, B. Z., Faller, J., and Neuper, C. (2012b). BCIs that use steady-state visual evoked potentials or slow cortical potentials. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Aloise, F., Lasorsa, I., Brouwer, A., Mattia, D., Babiloni, F., Salinari, S., Marciani, M., and Cincotti, F. (2007). Multimodal stimulation for a P300-based BCI. *International Journal of Bioelectromagnetism*, 9(3):128–130.

Bach, J. (1993). Amyotrophic lateral sclerosis. communication status and survival with ventilatory support. *American Journal of Physical Medicine & Rehabilitation / Association of Academic Physiatrists*, 72(6):343–349.

Bauer, G., Gerstenbrand, F., and Rumpl, E. (1979). Varieties of the locked-in syndrome. *Journal of Neurology*, 221(2):77–91.

Bentin, S., Allison, T., Puce, A., Perez, E., and McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8(6):551–565. PMID: 20740065 PMCID: PMC2927138.

Berntson, G. G., Bigger, J T, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., Nagaraja, H. N., Porges, S. W., Saul, J. P., Stone, P. H., and van der Molen, M. W. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*, 34(6):623–648. PMID: 9401419.

Berntson, G. G., Quigley, K. S., Jang, J. F., and Boysen, S. T. (1990). An approach to artifact identification: Application to heart period data. *Psychophysiology*, 27(5):586–598.

Berntson, G. G. and Stowell, J. R. (1998). ECG artifacts and heart period variability: don't miss a beat! *Psychophysiology*, 35(1):127–132. PMID: 9499713.

Bertsch, K., Hagemann, D., Naumann, E., Schächinger, H., and Schulz, A. (2012). Stability of heart rate variability indices reflecting parasympathetic activity. *Psychophysiology*, 49(5):672–682. PMID: 22335779.

Birbaumer, N. (2006). Breaking the silence: Brain-computer interfaces (BCI) for communication and motor control. *Psychophysiology*, 43(6):517–532.

Birbaumer, N. and Cohen, L. G. (2007). Brain-computer interfaces: communication and restoration of movement in paralysis. *The Journal of Physiology*, 579(Pt 3):621–636. PMID: 17234696.

Birbaumer, N., Elbert, T., Canavan, A., and Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. *Physiological Reviews*, 70(1):1–41.

Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., Perelmouter, J., Taub, E., and Flor, H. (1999). A spelling device for the paralysed. *Nature*, 398(6725):297–298. PMID: 10192330.

Birbaumer, N., Murguialday, A. R., and Cohen, L. (2008). Brain-computer interface in paralysis. *Current Opinion in Neurology*, 21(6):634–638. PMID: 18989104.

Birbaumer, N., Piccione, F., Silvoni, S., and Wildgruber, M. (2012). Ideomotor silence: the case of complete paralysis and brain-computer interfaces (BCI). *Psychological Research*, 76(2):183–191. PMID: 22252304.

Blankertz, B., Sannelli, C., Halder, S., Hammer, E. M., Kübler, A., Müller, K.-R., Curio, G., and Dickhaus, T. (2010). Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*, 51(4):1303–1309. PMID: 20303409.

Bogousslavsky, J., Melle, G. V., and Regli, F. (1988). The lausanne stroke registry: analysis of 1,000 consecutive patients with first stroke. *Stroke*, 19(9):1083–1092.

Borasio, G. D. and Voltz, R. (1997). Palliative care in amyotrophic lateral sclerosis. *Journal of Neurologyl*, 244(4):S11–S17.

Botrel, L., Acqualagna, L., Blankertz, B., Kaufmann, T., and Kübler, A. (2013). Relaxation exercise improves SMR-BCI performance. In *Procedings of the Fourth Workshop on Tools for Brain-Computer Interaction (TOBI)*, Sion, Switzerland.

Brouwer, A.-M. and van Erp, J. B. F. (2010). A tactile P300 brain-computer interface. *Frontiers in Neuroscience*, 4:19. PMID: 20582261.

Brouwer, A.-M., van Erp, J. B. F., Aloise, F., and Cincotti, F. (2010). Tactile, visual, and bimodal P300s: Could bimodal P300s boost BCI performance? *SRX Neuroscience*, 2010:1–9.

Brunner, P., Joshi, S., Briskin, S., Wolpaw, J. R., Bischof, H., and Schalk, G. (2010). Does the 'P300' speller depend on eye gaze? *Journal of Neural Engineering*, 7(5):056013. PMID: 20858924.

Bruno, M.-A., Schnakers, C., Damas, F., Pellas, F., Lutte, I., Bernheim, J., Majerus, S., Moonen, G., Goldman, S., and Laureys, S. (2009). Locked-in syndrome in children: Report of five cases and review of the literature. *Pediatric Neurology*, 41(4):237–246.

Byrne, S., Walsh, C., Lynch, C., Bede, P., Elamin, M., Kenna, K., McLaughlin, R., and Hardiman, O. (2011). Rate of familial amyotrophic lateral sclerosis: a systematic review and meta-analysis. *Journal of Neurology, Neurosurgery and Psychiatry*, 82(6):623–627.

Caplan, L. R., Wityk, R. J., Glass, T. A., Tapia, J., Pazdera, L., Chang, H.-M., Teal, P., Dashe, J. F., Chaves, C. J., Breen, J. C., Vemmos, K., Amarenco, P., Tettenborn, B., Leary, M., Estol, C., Dewitt, L. D., and Pessin, M. S. (2004). New england medical center posterior circulation registry. *Annals of Neurology*, 56(3):389–398.

Charcot, J. M. and Joffroy, A. (1869). Deux cas d'atrophie musculaire progressive avec lesions de la substance grise et des faisceaux anero-lateraux de la moelle epiniere. *Archives of Physiology, Neurology and Pathology*, 2(2):744.

Cifuentes-Diaz, C., Frugier, T., and Melki, J. (2002). Spinal muscular atrophy. *Seminars in Pediatric Neurology*, 9(2):145–150.

Cincotti, F., Kauhanen, L., Aloise, F., Palomäki, T., Caporusso, N., Jylänki, P., Mattia, D., Babiloni, F., Vanacker, G., Nuttin, M., Marciani, M. G., and Del R Millán, J. (2007). Vibrotactile feedback for brain-computer interface operation. *Computational Intelligence and Neuroscience*, page 48937. PMID: 18354734.

Cipresso, P., Meriggi, P., Carelli, L., Solca, F., Poletti, B., Lulé, D., Ludolph, A. C., Silani, V., and Riva, G. (2012). Brain computer interface and eye-tracking for neuropsychological assessment of executive functions: A pilot study. In *Computing Paradigms for Mental Health. Proceedings of the 2nd International Workshop on Computing Paradigms for Mental Health MindCare 2012*, pages 79–88, Vilamoura, Algarve, Portugal. SciTePress.

D8.12 (2013). Final report on dry electrodes. eu fp7-224631, tobi project, www.tobi-project.org.

Daucé, E. and Proix, T. (2013). P300-speller adaptivity to change with a backspace key. In *Procedings of the Fourth Workshop on Tools for Brain Computer Interaction (TOBI)*, Sion, Switzerland.

Debener, S., Minow, F., Emkes, R., Gandras, K., and de Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology*, 49(11):1617–1621.

Doble, J. E., Haig, A. J., Anderson, C., and Katz, R. (2003). Impairment, activity, participation, life satisfaction, and survival in persons with locked-in syndrome for over a decade: follow-up on a previously reported cohort. *The Journal of Head Trauma Rehabilitation*, 18(5):435–444. PMID: 12973273.

Dollfus, P., Milos, P. L., Chapuis, A., Real, P., Orenstein, M., and Soutter, J. W. (1990). The locked-in syndrome: a review and presentation of two chronic cases. *Paraplegia*, 28(1):5–16. PMID: 2152394.

Donoghue, J. (2012). BCIs that use signals recorded in motor cortex. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Eimer, M. (2000). Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology*, 111(4):694–705.

Eisen, A., Schulzer, M., MacNeil, M., Pant, B., and Mak, E. (1993). Duration of amyotrophic lateral sclerosis is age dependent. *Muscle and Nerve*, 16(1):27–32.

Fabiani, M., Gratton, G., and Federmeier, K. (2007). Event-related brain potentials: methods, theory and applications. In *Handbook of Psychophysiology*. Cambridge University Press, 3 edition.

Farwell, L. A. and Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6):510–523. PMID: 2461285.

Furdea, A., Halder, S., Krusienski, D. J., Bross, D., Nijboer, F., Birbaumer, N., and Kübler, A. (2009). An auditory oddball (P300) spelling system for brain-computer interfaces. *Psychophysiology*, 46(3):617–625. PMID: 19170946.

Galán, F., Nuttin, M., Lew, E., Ferrez, P., Vanacker, G., Philips, J., and Millán, J. d. R. (2008). A brain-actuated wheelchair: Asynchronous and non-invasive brain-computer interfaces for continuous control of robots. *Clinical Neurophysiology*, 119(9):2159–2169.

Glass, T., Hennessey, P., Pazdera, L., Chang, H., Wityk, R., Dewitt, L., Pessin, M., and LR, C. (2002). Outcome at 30 days in the new england medical center posterior circulation registry. *Archives of Neurology*, 59(3):369–376.

Grehl, H. and Reinhardt, F. (2008). Hirnstamminfarkt (Hirnstammsyndrome). In *Checkliste Neurologie*, pages 316–322. Thieme, 4th edition.

Grosse-Wentrup, M., Mattia, D., and Oweiss, K. (2011). Using brain-computer interfaces to induce neural plasticity and restore function. *Journal of Neural Engineering*, 8(2):025004.

Grossman, P. and Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*, 74(2):263–285. PMID: 17081672.

Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., Gramatica, F., and Edlinger, G. (2009). How many people are able to control a P300-based brain-computer interface (BCI)? *Neuroscience Letters*, 462(1):94–98. PMID: 19545601.

Halder, S., Agorastos, D., Veit, R., Hammer, E. M., Lee, S., Varkuti, B., Bogdan, M., Rosenstiel, W., Birbaumer, N., and Kübler, A. (2011). Neural mechanisms of brain-computer interface control. *NeuroImage*, 55(4):1779–1790. PMID: 21256234.

Halder, S., Hammer, E. M., Kleih, S. C., Bogdan, M., Rosenstiel, W., Birbaumer, N., and Kübler, A. (2013). Prediction of auditory and visual P300 brain-computer interface aptitude. *PLoS ONE*, 8(2):e53513.

Halder, S., Rea, M., Andreoni, R., Nijboer, F., Hammer, E. M., Kleih, S. C., Birbaumer, N., and Kübler, A. (2010). An auditory oddball brain-computer interface for binary choices. *Clinical Neurophysiology*, 121(4):516–523. PMID: 20093075.

Hammer, E. M., Halder, S., Blankertz, B., Sannelli, C., Dickhaus, T., Kleih, S., Müller, K.-R., and Kübler, A. (2012). Psychological predictors of SMR-BCI performance. *Biological Psychology*, 89(1):80–86. PMID: 21964375.

Hand, C. K. and Rouleau, G. A. (2002). Familial amyotrophic lateral sclerosis. *Muscle and Nerve*, 25(2):135–159.

Hardiman, O., Berg, L. H. v. d., and Kiernan, M. C. (2011). Clinical diagnosis and management of amyotrophic lateral sclerosis. *Nature Reviews Neurology*, 7(11):639–649.

Hedman, A. E., Tahvanainen, K. U. O., Hartikainen, J. E. K., and Hakumäki, M. O. K. (1995). Effect of sympathetic modualtion and sympathovagal interaction on heart rate variability in anaesthetized dogs. *Acta Physiologica Scandinavica*, 155(2):205–214.

Hill, N. and Schölkopf, B. (2012). An online brain-computer interface based on shifting attention to concurrent streams of auditory stimuli. *Journal of Neural Engineering*, 9(2):026011.

Hill, N. J., Lal, T. N., Bierig, K., Birbaumer, N., and Schölkopf, B. (2005). An auditory paradigm for brain–computer interfaces. In *Saul, L.K., Weiss, Y., Bottou, L.Advance in Neural Information Processing Systems*, volume 17, pages 569–576. MIT Press, MIT Press, Cambridge, MA.

Hinterberger, T., Schmidt, S., Neumann, N., Mellinger, J., Blankertz, B., Curio, G., and Birbaumer, N. (2004). Brain-computer communication and slow cortical potentials. *IEEE Transactions on Bio-Medical Engineering*, 51(6):1011–1018. PMID: 15188872.

Hoffmann, U., Vesin, J.-M., Ebrahimi, T., and Diserens, K. (2008). An efficient P300-based brain-computer interface for disabled subjects. *Journal of Neuroscience Methods*, 167(1):115–125.

Höhne, J., Schreuder, M., Blankertz, B., and Tangermann, M. (2010). Two-dimensional auditory P300 speller with predictive text system. *Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2010:4185–4188. PMID: 21096889.

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

Holz, E., Botrel, L., and Kübler, A. (2013). Bridging gaps: Long-term independent BCI home-use by a locked-in end-user. In *Procedings of the fourth workshop on tools for brain computer interaction (TOBI)*, Sion, Switzerland.

Holz, E., Höhne, J., Staiger-Sälzer, P., Tangermann, M., and Kübler, A. (subm.). BCI-controlled gaming: Evaluation of usability by severely motor restricted end-users. *Artificial Intelligence in Medicine*.

Holz, E. M., Kaufmann, T., Desideri, L., Malavasi, M., Hoogerwerf, E.-J., and Kübler, A. (2012). User centred design in BCI development. In Allison, B. Z., Dunne, S., Leeb, R., Millán, J. D. R., and Nijholt, A., editors, *Towards Practical Brain-Computer Interfaces*, Biological and Medical Physics, Biomedical Engineering, pages 155–172. Springer.

Hudson, A. J. (1981). Amyotrophic lateral sclerosis and its association with dementia, parkinsonism and other neurological disorders: A review. *Brain*, 104(2):217–247.

Huggins, J. and Zeitlin, D. (2012). BCI applications. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Iturrate, I., Antelis, J., Kübler, A., and Minguez, J. (2009). A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation. *IEEE Transactions on Robotics*, 25(3):614 –627.

Jin, J., Allison, B. Z., Kaufmann, T., Kübler, A., Zhang, Y., Wang, X., and Cichocki, A. (2012). The changing face of P300 BCIs: a comparison of stimulus changes in a P300 BCI involving faces, emotion, and movement. *PLoS ONE*, 7(11):e49688.

Jin, J., Allison, B. Z., Sellers, E. W., Brunner, C., Horki, P., Wang, X., and Neuper, C. (2011). An adaptive P300-based control system. *Journal of Neural Engineering*, 8(3):036006. PMID: 21474877.

Kasarskis, E. J., Berryman, S., Vanderleest, J. G., Schneider, A. R., and McClain, C. J. (1996). Nutritional status of patients with amyotrophic lateral sclerosis: relation to the proximity of death. *American Journal of Clinical Nutrition*, 63(1):130–137.

Käthner, I., Ruf, C. A., Pasqualotto, E., Braun, C., Birbaumer, N., and Halder, S. (2013). A portable auditory P300 brain-computer interface with directional cues. *Clinical Neurophysiology*, 124(2):327–338. PMID: 22959257.

Kato, S., Hayashi, H., and Yagishita, A. (1993). Involvement of the frontotemporal lobe and limbic system in amyotrophic lateral sclerosis: As assessed by serial computed tomography and magnetic resonance imaging. *Journal of the Neurological Sciences*, 116(1):52–58.

Kaufmann, T., Hammer, E. M., and Kübler, A. (2011a). ERPs contributing to classification in the ”P300” BCI. In *Procedings of the fifth International BCI Conference*, Graz, Austria.

Kaufmann, T., Holz, E., and Kübler, A. (2013a). The importance of user-centred design in BCI development: A case study with a locked-in patient. In *Procedings of the fourth Workshop on Tools for Brain Computer Interaction (TOBI)*, Sion, Switzerland.

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

Kaufmann, T., Schulz, S. M., Grünzinger, C., and Kübler, A. (2011b). Flashing characters with famous faces improves ERP-based brain-computer interface performance. *Journal of Neural Engineering*, 8(5):056016. PMID: 21934188.

Kaufmann, T., Schulz, S. M., Köblitz, A., Renner, G., Wessig, C., and Kübler, A. (2013b). Face stimuli effectively prevent brain-computer interface inefficiency in patients with neurodegenerative disease. *Clinical Neurophysiology*, 124(5):893–900. PMID: 23246415.

Kaufmann, T., Sütterlin, S., Schulz, S. M., and Vögele, C. (2011c). ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. *Behavior Research Methods*, 43(4):1161–1170.

Kaufmann, T., Vögele, C., Sütterlin, S., Lukito, S., and Kübler, A. (2012a). Effects of resting heart rate variability on performance in the P300 brain-computer interface. *International Journal of Psychophysiology*, 83(3):336–341.

Kaufmann, T., Völker, S., Gunesch, L., and Kübler, A. (2012b). Spelling is just a click away - a user-centered brain-computer interface including auto-calibration and predictive text entry. *Frontiers in Neuroscience*, 6:72. PMID: 22833713.

Kiernan, J. A. and Hudson, J. (1994). Frontal lobe atrophy in motor neuron diseases. *Brain*, 117(4):747–757.

Kiernan, M. C., Vucic, S., Cheah, B. C., Turner, M. R., Eisen, A., Hardiman, O., Burrell, J. R., and Zoing, M. C. (2011). Amyotrophic lateral sclerosis. *The Lancet*, 377(9769):942–955.

Kleih, S. C., Kaufmann, T., Zickler, C., Halder, S., Leotta, F., Cincotti, F., Aloise, F., Riccio, A., Herbert, C., Mattia, D., and Kübler, A. (2011). Out of the frying pan into the fire-the P300-based BCI faces real-world challenges. In *Progress in Brain Research*, volume 194, pages 27–46. Elsevier.

Kleih, S. C., Nijboer, F., Halder, S., and Kübler, A. (2010). Motivation modulates the P300 amplitude during brain-computer interface use. *Clinical Neurophysiology*, 121(7):1023–1031. PMID: 20188627.

Klobassa, D. S., Vaughan, T. M., Brunner, P., Schwartz, N. E., Wolpaw, J. R., Neuper, C., and Sellers, E. W. (2009). Toward a high-throughput auditory P300-based brain-computer interface. *Clinical Neurophysiology*, 120(7):1252–1261. PMID: 19574091.

Krusienski, D. J., McFarland, D. J., and Principe, J. C. (2012). BCI signal processing: feature extraction. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2006). A comparison of classification techniques for the P300 speller. *Journal of Neural Engineering*, 3(4):299–305.

Krusienski, D. J., Sellers, E. W., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2008). Toward enhanced P300 speller performance. *Journal of Neuroscience Methods*, 167(1):15–21. PMID: 17822777.

Kübler, A. and Birbaumer, N. (2008). Brain-computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? *Clinical Neurophysiology*, 119(11):2658–2666. PMID: 18824406 PMCID: PMC2644824.

Kübler, A., Blankertz, B., Müller, K.-R., and Neuper, C. (2011). A model of BCI-control. In *Procedings of the fifth International BCI Conference*, Graz, Austria.

Kübler, A., Furdea, A., Halder, S., Hammer, E. M., Nijboer, F., and 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(1):90–100.

Kübler, A., Holz, E. M., and Kaufmann, T. (2013). A user centred approach for bringing BCI controlled applications to end-users. In *Brain Computer Interface*. InTech. [in press].

Kübler, A., Kotchoubey, B., Hinterberger, T., Ghanayim, N., Perelmouter, J., Schauer, M., Fritsch, C., Taub, E., and Birbaumer, N. (1999). The thought translation device: a neurophysiological approach to communication in total motor paralysis. *Experimental Brain Research*, 124(2):223–232. PMID: 9928845.

Kübler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J. R., and Birbaumer, N. (2001a). Brain-computer communication: unlocking the locked in. *Psychological Bulletin*, 127(3):358–375. PMID: 11393301.

Kübler, A. and Müller, K.-R. (2007). An introduction to brain-computer interfacing. In *Dornhege, G., Millán, J. del. R., Hinterberger, T., McFarland, D.J., Müller, K.R., Toward brain-computer interfacing*, pages 1–25. Cambridge: MIT Press.

Kübler, A., Neumann, N., Kaiser, J., Kotchoubey, B., Hinterberger, T., and Birbaumer, N. P. (2001b). Brain-computer communication: Self-regulation of slow cortical potentials for verbal communication. *Archives of Physical Medicine and Rehabilitation*, 82(11):1533–1539.

Kübler, A., Neumann, N., Wilhelm, B., Hinterberger, T., and Birbaumer, N. (2004). Predictability of brain-computer communication. *Journal of Psychophysiology*, 18(2-3):121–129.

Kübler, A., Nijboer, F., Mellinger, J., Vaughan, T. M., Pawelzik, H., Schalk, G., McFarland, D. J., Birbaumer, N., and Wolpaw, J. R. (2005a). Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, 64(10):1775–1777.

Kübler, A., Winter, S., Ludolph, A. C., Hautzinger, M., and Birbaumer, N. (2005b). Severity of depressive symptoms and quality of life in patients with amyotrophic lateral sclerosis. *Neurorehabilitation and Neural Repair*, 19(3):182–193.

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

Lenhardt, A., Kaper, M., and Ritter, H. J. (2008). An adaptive P300-based online brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(2):121–130. PMID: 18403280.

Liu, T., Goldberg, L., Gao, S., and Hong, B. (2010). An online brain-computer interface using non-flashing visual evoked potentials. *Journal of Neural Engineering*, 7(3):036003.

Logroscino, G., Traynor, B. J., Hardiman, O., Chio, A., Mitchell, D., Swingler, R. J., Millul, A., Benn, E., and Beghi, E. (2010). Incidence of amyotrophic lateral sclerosis in europe. *Journal of Neurology, Neurosurgery and Psychiatry*, 81(4):385–390.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Mit Press.

Ludolph, A. C., Brettschneider, J., and Weishaupt, J. H. (2012). Amyotrophic lateral sclerosis. *Current Opinion in Neurology*, 25(5):530–535. PMID: 22918486.

Lulé, D., Häcker, S., Ludolph, A., Birbaumer, N., and Kübler, A. (2008). Depression and quality of life in patients with amyotrophic lateral sclerosis. *Deutsches Ärzteblatt international*, 105(23):397–403. PMID: 19626161.

Mak, J. N., Arbel, Y., Minett, J. W., McCane, L. M., Yuksel, B., Ryan, D., Thompson, D., Bianchi, L., and Erdogmus, D. (2011). Optimizing the P300-based brain-computer interface: current status, limitations and future directions. *Journal of Neural Engineering*, 8(2):025003. PMID: 21436525.

Martens, S. M. M., Hill, N. J., Farquhar, J., and Schölkopf, B. (2009). Overlap and refractory effects in a brain-computer interface speller based on the visual P300 event-related potential. *Journal of Neural Engineering*, 6(2):026003.

Martinez, P., Bakardjian, H., and Cichocki, A. (2007). Fully online multicommand brain-computer interface with visual neurofeedback using SSVEP paradigm. *Computational Intelligence and Neuroscience*, page 94561. PMID: 18354725.

Mazzini, L., Corra, T., Zaccala, M., Mora, G., Del Piano, M., and Galante, M. (1995). Percutaneous endoscopic gastrostomy and enteral nutrition in amyotrophic lateral sclerosis. *Journal of Neurology*, 242(10):695–698. PMID: 8568533.

McFarland, D. J. and Krusienski, D. J. (2012). BCI signal processing: feature extraction. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

McGuire, V., Longstreth, W T, J., Koepsell, T. D., and van Belle, G. (1996). Incidence of amyotrophic lateral sclerosis in three counties in western washington state. *Neurology*, 47(2):571–573. PMID: 8757041.

Millán, J., Renkens, F., Mourino, J., and Gerstner, W. (2004). Noninvasive brain-actuated control of a mobile robot by human EEG. *IEEE Transactions on Biomedical Engineering*, 51(6):1026 –1033.

Millán, J. d. R. and Mouriño, J. (2003). Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):159–161. PMID: 12899262.

Mitchell, J. and Borasio, G. (2007). Amyotrophic lateral sclerosis. *The Lancet*, 369(9578):2031–2041.

Müller, K.-R. and Blankertz, B. (2006). Toward non-invasive brain-computer interfaces. *IEEE Signal Processing Magazine*, 23(5):125–128.

Müller-Putz, G. R., Scherer, R., Pfurtscheller, G., and Rupp, R. (2005). EEG-based neuroprosthesis control: a step towards clinical practice. *Neuroscience Letters*, 382(1):169–174.

Munsat, T. (1991). Workshop report. international SMA collaboration. *Neuromuscular Disorders*, 1(2):81.

Münssinger, J. I., Halder, S., Kleih, S. C., Furdea, A., Raco, V., Hösle, A., and Kübler, A. (2010). Brain painting: First evaluation of a new Brain-Computer interface application with ALS-Patients and healthy volunteers. *Frontiers in Neuroscience*, 4.

Neuper, C., Müller, G. R., Kübler, A., Birbaumer, N., and Pfurtscheller, G. (2003). Clinical application of an EEG-based brain-computer interface: a case study in a patient with severe motor impairment. *Clinical Neurophysiology*, 114(3):399–409. PMID: 12705420.

Neuper, C., Scherer, R., Reiner, M., and Pfurtscheller, G. (2005). Imagery of motor actions: differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG. *Brain Research. Cognitive Brain Research*, 25(3):668–677. PMID: 16236487.

Nijboer, F., Birbaumer, N., and Kübler, 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. PMID: 20700521.

Nijboer, F., Furdea, A., Gunst, I., Mellinger, J., McFarland, D. J., Birbaumer, N., and Kübler, A. (2008a). An auditory brain-computer interface (BCI). *Journal of Neuroscience Methods*, 167(1):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., Wolpaw, J. R., Birbaumer, N., and Kübler, A. (2008b). A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. *Clinical Neurophysiology*, 119(8):1909–1916. PMID: 18571984.

Niskanen, J.-P., Tarvainen, M. P., Ranta-aho, P. O., and Karjalainen, P. A. (2004). Software for advanced HRV analysis. *Computer Methods and Programs in Biomedicine*, 76(1):73–81.

Nunan, D., Donovan, G., Jakovljevic, D. G., Hodges, L. D., Sandercock, G. R. H., and Brodie, D. A. (2009). Validity and reliability of short-term heart-rate variability from the polar s810. *Medicine and Science in Sports and Exercise*, 41(1):243–250. PMID: 19092682.

Ogawa, S., Lee, T. M., Kay, A. R., and Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, 87(24):9868–9872. PMID: 2124706.

Otto, K., Ludwig, K. A., and Kipke, D. (2012). Acquiring brain signals from within the brain. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Pearn, J. (1978). Incidence, prevalence, and gene frequency studies of chronic childhood spinal muscular atrophy. *Journal of Medical Genetics*, 15(6):409–413.

Pfurtscheller, G., Guger, C., Müller, G., Krausz, G., and Neuper, C. (2000). Brain oscillations control hand orthosis in a tetraplegic. *Neuroscience Letters*, 292(3):211–214.

Pfurtscheller, G. and McFarland, D. (2012). BCIs that use sensorimotor rhythms. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Pfurtscheller, G., Müller, G. R., Pfurtscheller, J., Gerner, H. J., and Rupp, R. (2003). 'Thought'–control of functional electrical stimulation to restore hand grasp in a patient with tetraplegia. *Neuroscience Letters*, 351(1):33–36. PMID: 14550907.

Pfurtscheller, G. and Neuper, C. (1997). Motor imagery activates primary sensorimotor area in humans. *Neuroscience Letters*, 239(2-3):65–68. PMID: 9469657.

Pfurtscheller, G., Neuper, C., Flotzinger, D., and Pregenzer, M. (1997). EEG-based discrimination between imagination of right and left hand movement. *Electroencephalography and Clinical Neurophysiology*, 103(6):642–651. PMID: 9546492.

Piccione, F., Giorgi, F., Tonin, P., Priftis, K., Giove, S., Silvoni, S., Palmas, G., and Beverina, F. (2006). P300-based brain computer interface: Reliability and performance in healthy and paralysed participants. *Clinical Neurophysiology*, 117(3):531–537.

Pichiorri, F., De Vico Fallani, F., Cincotti, F., Babiloni, F., Molinari, M., Kleih, S. C., Neuper, C., Kübler, A., and Mattia, D. (2011). Sensorimotor rhythm-based brain-computer interface training: the impact on motor cortical responsiveness. *Journal of Neural Engineering*, 8(2):025020. PMID: 21436514.

Pires, G., Castelo-Branco, M., and Nunes, U. (2008). Visual P300-based BCI to steer a wheelchair: A bayesian approach. In *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008. EMBS 2008*, pages 658–661.

Pizzagalli, D. (2007). Electroencephalography and high-density electrophysiological source localization. In *Handbook of Psychophysiology*. Cambridge University Press, 3 edition.

Plum, F. and Posner, J. (1966). The diagnosis of stupor and coma. *Philadelphia, PA, FA Davis*.

Plum, F. and Posner, J. B. (1982). *The Diagnosis Of Stupor & Coma*. Oxford University Press.

Polich, J. (2007). Updating P300: An integrative theory of p3a and p3b. *Clinical Neurophysiology*, 118(10):2128–2148.

Porges, S. (1992). Autonomic regulation and attention. In *Attention and Information Processing in Infants and Adults: Perspectives from Human and Animal Research*, pages 201–224. Routledge.

Radunovic, A., Annane, D., Jewitt, K., and Mustfa, N. (2009). Mechanical ventilation for amyotrophic lateral sclerosis/motor neuron disease. *Cochrane Database of Systematic Reviews*, 4:CD004427. PMID: 19821325.

Ramsey, N. (2012). Signals reflecting brain metabolic activity. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Rebsamen, B., Guan, C., Zhang, H., Wang, C., Teo, C., Ang, M., and Burdet, E. (2010). A brain controlled wheelchair to navigate in familiar environments. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(6):590–598.

Riccio, A., Mattia, D., Simione, L., Olivetti, M., and Cincotti, F. (2012). Eye-gaze independent EEG-based brain-computer interfaces for communication. *Journal of Neural Engineering*, 9(4):045001.

Rowland, L. P. and Shneider, N. A. (2001). Amyotrophic lateral sclerosis. *New England Journal of Medicine*, 344(22):1688–1700. PMID: 11386269.

Ryan, D. B., Frye, G. E., Townsend, G., Berry, D. R., Mesa-G, S., Gates, N. A., and Sellers, E. W. (2011). Predictive spelling with a P300-based brain-computer interface: Increasing the rate of communication. *International Journal of Human-Computer Interaction*, 27(1):69–84. PMID: 21278858.

Salvaris, M. and Sepulveda, F. (2009). Visual modifications on the P300 speller BCI paradigm. *Journal of Neural Engineering*, 6(4):046011. PMID: 19602731.

Saul, J. P. (1990). Beat-to-beat variations of heart rate reflect modulation of cardiac autonomic outflow. *Physiology*, 5(1):32–37.

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., and Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Bio-Medical Engineering*, 51(6):1034–1043. PMID: 15188875.

Schreiber, H., Gaigalat, T., Wiedemuth-Catrinescu, U., Graf, M., Uttner, I., Muche, R., and Ludolph, A. C. (2005). Cognitive function in bulbar- and spinal-onset amyotrophic lateral sclerosis. a longitudinal study in 52 patients. *Journal of Neurology*, 252(7):772–781. PMID: 15742104.

Schreuder, M., Blankertz, B., and Tangermann, M. (2010). A new auditory multi-class brain-computer interface paradigm: spatial hearing as an informative cue. *PloS one*, 5(4):e9813. PMID: 20368976.

Schreuder, M., Hohne, J., Treder, M., Blankertz, B., and Tangermann, M. (2011a). Performance optimization of ERP-based BCIs using dynamic stopping. *Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2011:4580–4583. PMID: 22255357.

Schreuder, M., Rost, T., and Tangermann, M. (2011b). Listen, you are writing! speeding up online spelling with a dynamic auditory BCI. *Frontiers in Neuroscience*, 5:112. PMID: 22016719.

Schreuder, M., Thurlings, M. E., Brouwer, A.-M., Van Erp, J. B. F., and Tangermann, M. (2012). Exploring the use of tactile feedback in an ERP-based auditory BCI. *Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2012:6707–6710. PMID: 23367468.

Sellers, E. W., Arbel, Y., and Donchin, E. (2012). BCIs that use P300 event-related potentials. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Sellers, E. W. and Donchin, E. (2006). A P300-based brain-computer interface: initial tests by ALS patients. *Clinical Neurophysiology*, 117(3):538–548. PMID: 16461003.

Sellers, E. W., Krusienski, D. J., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2006). A P300 event-related potential brain-computer interface (BCI): the effects of matrix size and inter stimulus interval on performance. *Biological Psychology*, 73(3):242–252. PMID: 16860920.

Sellers, E. W., Vaughan, T. M., and Wolpaw, J. R. (2010). A brain-computer interface for long-term independent home use. *Amyotrophic Lateral Sclerosis*, 11(5):449–455.

Silvoni, S., Cavinato, M., Volpato, C., Ruf, C. A., Birbaumer, N., and Piccione, F. (2013). Amyotrophic lateral sclerosis progression and stability of brain-computer interface communication. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. PMID: 23445258.

Sitaram, R., Lee, S., and Birbaumer, N. (2012). BCIs that use brain metabolic signals. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Smith, E. and Delargy, M. (2005). Locked-in syndrome. *BMJ (Clinical Research ed.)*, 330(7488):406–409.

Srinivasan, R. (2012). Acquiring brain signals from outside the brain. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Stinear, C. M., Byblow, W. D., Steyvers, M., Levin, O., and Swinnen, S. P. (2006). Kinesthetic, but not visual, motor imagery modulates corticomotor excitability. *Experimental Brain Research*, 168(1-2):157–164.

Sutton, S., Braren, M., Zubin, J., and John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, 150(3700):1187–1188.

Talbot, K. (2009). Motor neuron disease: the bare essentials. *Practical Neurology*, 9(5):303–309. PMID: 19762894.

Talbot, K. and Davies, K. E. (2001). Spinal muscular atrophy. *Seminars in Neurology*, 21(02):189–198.

Tarvainen, M. P., Niskanen, J.-P., Lipponen, J. A., Ranta-aho, P. O., and Karjalainen, P. A. (2009). Kubios HRV - a software for advanced heart rate variability analysis. In Sloten, J. V., Verdonck, P., Nyssen, M., and Haueisen, J., editors, *4th European Conference of the International Federation for Medical and Biological Engineering*, number 22 in IFMBE Proceedings, pages 1022–1025. Springer Berlin Heidelberg.

TaskForce (1996). Heart rate variability. standards of measurement, physiological interpretation, and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology. *European Heart Journal*, 17(3):354–381. PMID: 8737210.

Thayer, J. and Siegle, G. (2002). Neurovisceral integration in cardiac and emotional regulation. *IEEE Engineering in Medicine and Biology Magazine*, 21(4):24–29.

Thayer, J. F. and Brosschot, J. F. (2005). Psychosomatics and psychopathology: looking up and down from the brain. *Psychoneuroendocrinology*, 30(10):1050–1058.

Thayer, J. F. and Lane, R. D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of Affective Disorders*, 61(3):201–216.

Thurlings, M. E., van Erp, J. B., Brouwer, A.-M., Blankertz, B., and Werkhoven, P. (2012). Control-display mapping in brain-computer interfaces. *Ergonomics*, 55(5):564–580. PMID: 22506977.

Touryan, J., Gibson, L., Horne, J. H., and Weber, P. (2011). Real-time measurement of face recognition in rapid serial visual presentation. *Frontiers in Psychology*, 2:42. PMID: 21716601.

Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C. K., Schwartz, N. E., Vaughan, T. M., Wolpaw, J. R., and Sellers, E. W. (2010). A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clinical Neurophysiology*, 121(7):1109–1120. PMID: 20347387.

Treder, M. S. and Blankertz, B. (2010). (c)overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and Brain Functions: BBF*, 6:28. PMID: 20509913.

van der Waal, M., Severens, M., Geuze, J., and Desain, P. (2012). Introducing the tactile speller: an ERP-based brain-computer interface for communication. *Journal of Neural Engineering*, 9(4):045002.

van Gerven, M., Farquhar, J., Schaefer, R., Vlek, R., Geuze, J., Nijholt, A., Ramsey, N., Haselager, P., Vuurpijl, L., Gielen, S., and Desain, P. (2009). The brain-computer interface cycle. *Journal of Neural Engineering*, 6(4):041001.

Vaughan, T. M., Sellers, E., and Wolpaw, J. R. (2012). Clinical evaluation of BCIs. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Vidal, J. J. (1973). Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering*, 2:157–180. PMID: 4583653.

Wilson, J., Guger, C., and Schalk, G. (2012). BCI hardware and software. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Wolpaw, J. R. and Wolpaw, E. W. (2012a). *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Wolpaw, J. R. and Wolpaw, E. W. (2012b). Brain-computer interfaces: something new under the sun. In *Wolpaw and Wolpaw, Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 1 edition.

Zerres, K. and Rudnik-Schöneborn, S. (1995). Natural history in proximal spinal muscular atrophy: Clinical analysis of 445 patients and suggestions for a modification of existing classifications. *Archives in Neurology*, 52(5):518–523.

Zhang, H., Guan, C., and Wang, C. (2008). Asynchronous P300-based brain-computer interfaces: a computational approach with statistical models. *IEEE Transactions on Biomedical Engineering*, 55(6):1754–1763. PMID: 18714840.

Zhang, Y., Zhao, Q., Jin, J., Wang, X., and Cichocki, A. (2012). A novel BCI based on ERP components sensitive to configural processing of human faces. *Journal of Neural Engineering*, 9(2):026018. PMID: 22414683.

Zickler, C., Di Donna, V., Kaiser, V., Al-Khodairy, A., Kleih, S. C., and Kübler, A. (2009). Bci applications for people with disabilities: defining user needs and user requirements. In *Emiliani, P.L., Burzagli, L.,Como, A., Gabbanini, F. and Salminen, A.-L. Assistive Technology from Adapted Equipment to Inclusive Environments, AAATE. 25 Assistive Technology Research Series*, pages 185–189. Amsterdam: IOS Press.

Zickler, C., Riccio, A., Leotta, F., Hillian-Tress, S., Halder, S., Holz, E., Staiger-Sälzer, P., Hoogerwerf, E.-J., Desideri, L., Mattia, D., and Kübler, A. (2011). A brain-computer interface as input channel for a standard assistive technology software. *Clinical EEG and Neuroscience*, 42(4):236–244.

Appendix A: Affidavit (Eidesstattliche Erklärung)

I hereby confirm that my thesis entitled *Brain-Computer Interfaces based on Event-Related Potentials: toward fast, reliable and easy-to-use communication systems for people with neurodegenerative disease* 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.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

....	....
Place, Date	Signature

Appendix B: Full reference to publications that were a result of this thesis

2013

- **Kaufmann, T.**, Schulz, S. M., Köblitz, A., Renner, G., Wessig, C., and Kübler, A. (2013). Face stimuli effectively prevent brain-computer interface inefficiency in patients with neurodegenerative disease. Clinical Neurophysiology, 124(5), 893-900, doi:10.1016/j.clinph.2012.11.006.

2012

- **Kaufmann, T.**, Völker, S., Gunesch, L., and Kübler, A. (2012). Spelling is just a click away - a user-centered brain-computer interface including auto-calibration and predictive text entry. Frontiers in Neuroscience, 6, 72, doi:10.3389/fnins.2012.00072.
- **Kaufmann, T.**, Vögele, C., Sütterlin, S., Lukito, S., and Kübler, A. (2012). Effects of resting heart rate variability on performance in the P300 brain-computer interface. International Journal of Psychophysiology, 83(3), 336-341, doi:10.1016/j.ijpsycho.2011.11.018.

2011

- **Kaufmann, T.**, Schulz, S. M., Grünzinger, C., and Kübler, A. (2011). Flashing characters with famous faces improves ERP-based brain-computer interface performance. Journal of Neural Engineering, 8(5), 056016. doi:10.1088/1741-2560/8/5/056016.
- **Kaufmann, T.**, Sütterlin, S., Schulz, S. M., and Vögele, C. (2011). ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. Behavior Research Methods, 43(4), 1161-1170, doi:10.3758/s13428-011-0107-7.

Appendix C: Approval of a ”Dissertation Based on Several Published Manuscripts”

Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

Graduate School
Life Sciences

Approval of a "Dissertation Based on Several Published Manuscripts"

for the doctoral researcher

Dipl.-Biol. Tobias Kaufmann

(Name)

who has accomplished a publication record significantly above average as documented in the attachment.

The **Section Speakers and the Thesis Committee** therefore approve a "Dissertation Based on Several Published Manuscripts".

The **Thesis Committee** additionally confirms that the doctoral researcher has fulfilled all requirements of the GSLS program "life science".

Thesis Committee

Supervisor	Name	Date	Signature
1	Prof. Dr. A. Kübler	13.12.2012	
2	Prof. Dr. W. Rössler	17.12.2012	
3	Prof. Dr. K. Schilling	13.12.2012	Schilling
4 (if applicable)			

Section Speakers

Speaker	Name	Date	Signature
1	Prof. Dr. Paul Pauli	13.02.2013	Prof. Dr. P. Pauli
2	Prof. Dr. E. Asan	22.02.2013	
3 (if applicable)	Prof. Dr. M. Sendtner	20.03.2013	

Appendix D: Statement on individual author contributions

“Dissertation Based on Several Published Manuscripts“

Statement on individual author contributions and on legal second publication rights

(if required use more sheets of this form)

Publication: Kaufmann, T, Sütterlin, S., Schulz, S. M., and Vögele, C. (2011). ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. Behavior Research Methods, 43(4), 1161-1170.

Participated in	**Author-Initials,** Responsibility decreasing from left to right				
Study Design	**TK**	SS	CV		
Data Collection	**TK**				
Data-Analysis and Interpretation	**TK**	SS	SMS		
Manuscript Writing	**TK**	SS	SMS	CV	

Explanations (if applicable): This is a software publication. Implementation of the software was completely done by **TK**

Publication: Kaufmann, T, Schulz, S. M., Grünzinger, C., and Kübler, A. (2011). Flashing characters with famous faces improves ERP-based brain-computer interface performance. Journal of Neural Engineering, 8(5), 056016.

Participated in	**Author-Initials,** Responsibility decreasing from left to right				
Study Design	SMS	**TK**	AK		
Data Collection	CG	**TK**			
Data-Analysis and Interpretation	**TK**	SMS	AK		
Manuscript Writing	**TK**	SMS	AK		

Explanations (if applicable):

Publication: Kaufmann, T, Vögele, C., Sütterlin, S., Lukito, S., and Kübler, A. (2012). Effects of resting heart rate variability on performance in the P300 brain-computer interface. International Journal of Psychophysiology, 83(3), 336-341.

Participated in	**Author-Initials,** Responsibility decreasing from left to right				
Study Design	CV	**TK**	AK	SS	
Data Collection	**TK**	SS	SL		
Data-Analysis and Interpretation	**TK**	SS	AK	CV	
Manuscript Writing	**TK**	CV	SS	AK	

Explanations (if applicable):

Publication: Kaufmann, T, Völker, S., Gunesch, L., and Kübler, A. (2012). Spelling is just a click away - a user-centered brain-computer interface including auto-calibration and predictive text entry. Frontiers in Neuroprosthetics, 6, 72.

Participated in	**Author-Initials,** Responsibility decreasing from left to right				
Study Design	**TK**	SV	AK		
Data Collection	LG	**TK**			
Data-Analysis and Interpretation	**TK**	AK			
Manuscript Writing	**TK**	AK			

Explanations (if applicable):

Publication: Kaufmann, T, Schulz, S. M., Köblitz, A., Renner, G., Wessig, C., and Kübler, A. (2013). Face stimuli effectively prevent brain-computer interface inefficiency in patients with neurodegenerative disease. Clinical Neurophysiology, 124(5), 893-900.

Participated in	**Author-Initials,** Responsibility decreasing from left to right				
Study Design	**TK**	SMS	AKü		
Data Collection	**TK**	AKö			
Data-Analysis and Interpretation	**TK**	SMS	AKü		
Manuscript Writing	**TK**	SMS	AKü	GR	CW

Explanations (if applicable):

I confirm that I have obtained permission from both the publishers and the co-authors for legal second publication.

I also confirm my primary supervisor's acceptance.

Tobias Kaufmann Würzburg

Doctoral Researcher's Name Date Place Signature

Appendix E: Curiculum vitae

CURRICULUM VITAE

Tobias Kaufmann

Department of Psychology I, University of Würzburg

EDUCATION & RESEARCH EXPERIENCE

Since 2010 **Doctoral student at Department of Psychology I, University of Würzburg**
Thesis title: "*Brain computer interfaces based on event-related potentials: toward fast, reliable and easy-to-use communication systems for people with neurodegenerative disease*". To be submitted to the Graduate School of Life Science.

03-04/2010 Research project at Graz University of Technology, Graz, Austria

2004 – 2009 **Studies of Biology at Albert Ludwig University of Freiburg**
- Neurobiology and Biophysics (main subject)
- Neurophysiology (complementary)
- Bioinformatics (complementary)
- Computer Science (complementary)

2009 Graduation in Biology (Diplom) with the thesis "*Structural motor learning of dynamic and visuomotor transformations*" at the Institute for Biology I, Freiburg, Germany

2008 Internship at the Clinical & Health Psychology Research Centre at Roehampton University, London, United Kingdom

2008 Participation at the 4th Summer School on Emerging Technologies in Biomedicine, University of Patras, Greece

2007 – 2008 Research assistance
- Bernstein Center for Computational Neuroscience, Freiburg, Germany
- Institute for Biology I, Freiburg, Germany

PROFESSIONAL EXPERIENCE

Since 2010 **Supervision at the University of Würzburg**
Supervision and guide of 4 diploma students (6-9 month each), one master student (9 month) and 3 internships (several months each) in the fields of Psychology, Biology and Medical Engineering

Since 2011 **Reviewer for Scientific Journals and Conferences**
- *Journals:* Journal of Neural Engineering, Psychophysiology, NeuroComputing, IEEE Trans. on Biomedical Engineering, Journal of Neuroscience Methods
- *Conferences*: 5th IEEE Conference on Neural Engineering, 5th International Brain Computer Interface Meeting, BBCI Workshop 2012, 3rd and 4th TOBI Workshop

Jan. 17, 2013 **Invited talk at the Cognitive Science Lab, University of Osnabrück**
Title of talk: Toward brain computer interface based wheelchair control: Utilizing event-related potentials to deliver navigation commands

Spring 2012 **Organization of the 3rd Workshop on Brain Computer Interaction (TOBI)**

AWARDS

Feb. 8, 2013 **Best Publication Award** (includes prize money of 500€) of the Research Training Group RTG1253/1 (Deutsche Forschungs Gesellschaft) for the publication *Kaufmann, T., Schulz, S. M., Köblitz, A., Renner, G., Wessig, C., & Kübler, A. (2013). Face stimuli effectively prevent brain-computer interface inefficiency in patients with neurodegenerative disease. Clinical Neurophysiology,* 124(5), 893-900.

March 21, 2012 **Best Talk Award** (includes prize money of 300 €) for the talk *ERP-BCI based communication using a straightforward, user-centered software tool including auto-calibration and predictive text entry*, 3rd Workshop on Tools for Brain Computer Interaction (TOBI), Würzburg

GRANTS

May 19, 2008 **Travel Grant of the European Cooperation in the field of Scientific and Technical Research (COST)**

PUBLICATIONS

Research articles:

2013

- **Kaufmann, T.**, Schulz, S. M., Köblitz, A., Renner, G., Wessig, C., & Kübler, A. (2013). Face stimuli effectively prevent brain-computer interface inefficiency in patients with neurodegenerative disease. *Clinical Neurophysiology,* 124(5), 893-900. doi:10.1016/j.clinph.2012.11.006.

2012

- Jin, J., Allison, B. Z., **Kaufmann, T.**, Kübler, A., Zhang, Y., Wang, X., & Cichocki, A. (2012). The Changing Face of P300 BCIs: A Comparison of Stimulus Changes in a P300 BCI Involving Faces, Emotion, and Movement. *PloS one*, 7(11), e49688. doi:10.1371/journal.pone.0049688.
- **Kaufmann, T.**, Völker S., Gunesch L. & Kübler A. (2012). Spelling is just a click away – a user-centered brain-computer interface including auto-calibration and predictive text entry. *Frontiers in Neuroscience,* **6**:72. doi: 10.3389/fnins.2012.00072.
- **Kaufmann, T.**, Vögele, C., Sütterlin, S., Lukito, S., & Kübler, A. (2012). Effects of resting heart rate variability on performance in the P300 brain-computer interface. *International Journal of Psychophysiology*, 83(3), 336–341. doi:10.1016/j.ijpsycho.2011.11.018.

2011

- **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. doi:10.1088/1741-2560/8/5/056016.
- **Kaufmann, T.**, Sütterlin, S., Schulz, S.M., & Vögele, C. (2011). ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. *Behavior Research Methods,* 43(4), 1161–1170. doi:10.3758/s13428-011-0107-7.
- Kleih, S. C., **Kaufmann, T**., Zickler, C., Halder, S., Leotta, F., Cincotti, F., Aloise, F., Riccio, A., Herbert, C., Mattia, D., & Kübler, A. (2011). Out of the frying pan into the fire-the P300-based BCI faces real-world challenges. Progress in Brain Research, 194, 27-46. doi:10.1016/B978-0-444-53815-4.00019-4.

Book chapters:

In press

- Kübler, A., Holz, E., & **Kaufmann, T.** (in press). A user centred approach for bringing BCI controlled applications to end-users. In Brain-Computer Interface, ISBN 980-953-307-960-3.

2012

- Holz, E., **Kaufmann, T.**, Desidiri, L., Malavasi, M., Hoogerwerf, E.J., & Kübler A. (2012) User Centred Design in BCI Development, In: Allison, B., Dunne, S., Leeb, R., Millán, J.d.R., Nijholt, A. (Eds), Towards Practical Brain-Computer Interfaces: Bridging the Gap from Research to Real-World Applications. Series: Biological and Medical Physics, Biomedical Engineering: Springer-Verlag GmbH Berlin, 2012 http://www.springerlink.com/content/h345413430217121/.

Conference contributions (first author only):

Conference talks:

- **Kaufmann, T.**, Völker, S., Gunesch, L., & Kübler, A. (2012), ERP-BCI based communication using a straightforward, user-centered software tool including auto-calibration and predictive text entry, *3rd Workshop on Tools for Brain Computer Interaction (TOBI), Würzburg, March 21, 2012.*
- **Kaufmann, T.**, Holz, E., & Kübler, A. (2013). The importance of user-centred design in BCI development: A case study with a locked-in patient. *4th Workshop on Tools for Brain Computer Interaction (TOBI), Sion, January 17, 2013.*

Poster presentations:

- **Kaufmann, T.**, Schulz, S.M., Köblitz, A., Renner, G., Wessig, C., & Kübler, A., (2013). Face stimuli prevent ERP-BCI inefficiency in users with neurodegenerative disease. *4th Workshop on Tools for Brain Computer Interaction (TOBI), Sion, January 17, 2013.*
- **Kaufmann, T**., Schulz, S. M., & Kübler, A. (2012). Flashing Characters with Famous Faces Significantly Improves Online Bit Rate of ERP-BCIs. *3rd Workshop on Tools for Brain Computer Interaction (TOBI)*, University of Würzburg, Würzburg, Gernany.
- **Kaufmann, T.**, Hammer, E.M., & Kübler, A. (2011), ERPs contributing to classification in the "P300"-BCI, *Proceedings of the 5th International Brain-Computer Interface Conference*, 136-139.
- **Kaufmann, T.**, Williamson, J., Hammer, E., Murray-Smith, R. & Kübler, A. (2010), Visually multimodal vs. classic unimodal feedback approach for SMR-BCIs: A comparison study, *2nd Workshop on Tools for Brain Computer Interaction (TOBI)*, Fondazione Santa Lucia, Rome, Italy.
- **Kaufmann, T.**, Vögele, C., Sütterlin, S., Lukito, S. & Kübler, A. (2010), Cardiac autonomic control as physiological trait predicts performance in a P300 BCI, *4th International BCI Meeting*, Monterey, California, USA.
- **Kaufmann, T.**, Vögele, C., Sütterlin, S., Lukito, S. & Kübler, A. (2010), Effects of cardiac autonomic balance on performance in a P300 brain computer interface, *1st Workshop on Tools for Brain Computer Interaction (TOBI)*, University of Technology, Graz, Austria.

------------------- ---------------------------------------

Date Signature