UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Sensorimotor rhythm brain-computer interface – A game-based online co-adaptive training

José Diogo Marques da Silva Branco da Cunha

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil de Engenharia Clínica e Instrumentação Médica

Dissertação orientada por:
Prof. Reinhold Scherer
Prof. Alexandre Andrade

2018

Acknowledgements

The last semester was surely one of the most eventful of my life. I share the memory of each episode of it with different people and groups to whom I must thank dearly.

First, I must acknowledge Reini for supervising, and supporting the whole project, for always showing me the bigger picture in my academic and personal work, and for, alongside with it, becoming a dear personal friend whose wisdom I sought frequently.

I must acknowledge Professor Alexandre Andrade for the solicitude and wisdom of his advice.

I would like to mention the Institute of Neural Engineering group for the welcoming spirit, in particular, I want to thank Florian Hubmann for the tireless help during the project and for the immense kindness and enthusiasm that made my stay so much better.

I would like to express my deep appreciation to my parents who supported me throughout my studies and always pushed me to expand my horizons and go further, both at home and abroad.

Thank you Inês, for being an unweariable source of serenity and love, whose advice made every aspect of my stay better.

Finally, I would like to show my appreciation to the support of the Erasmus+ program of the European Union for my internship abroad.

Abstract

Brain-Computer Interface (BCI) technology translates brain signals into messages. BCI users are thus enabled to interact with the environment by thought, or more generally speaking by mental processes. Event-related desynchronization (ERD) based BCIs use the detection of changes in the spontaneous electroencephalogram (EEG) signal. Different mental processes induce power decreases (ERD) or increases (event-related synchronization, ERS) in different frequencies and different areas of the brain. These differences can be measured and classified. Operating a non-invasive EEG based sensorimotor rhythm BCI is a skill that typically requires extensive training. Lately, online co-adaptive feedback training approaches achieved promising results after short periods of training. Does this also mean that users can have meaningful BCI-based interactions after training, when the BCI is no longer adapting, like in a real-life scenario? To answer this question an online study was conducted with 20 naïve (first time) users. After a short (less than 20 minutes) setup, the users trained to gain BCI control by playing a Whack-A-Mole game where they would have to perform Motor Imagery (imagination of a specific movement - MI) to control a hammer to hit a mole. The game was played for about 30 minutes. During this time, the user learns to perform MI with online feedback from the game and the BCI parameters recurrently adapt to the user's EEG patterns every ~1 minute. This recurrent adaptation allows different users to use slightly different strategies and produce ERDs in different frequencies and brain areas without loss of performance. After 30 minutes of training the adaptation was stopped and the users continued playing the game with the trained BCI for another 20 minutes. The BCI parameters were calibrated with data from the adaptive stage and kept fixed in the last 20 minutes. Our hypothesis is that once a system was co-adaptively trained it can maintain its performance without recurrent adaptation. Eighteen out of the twenty users were able to control the BCI and play the game. Seventeen out of the eighteen were able to improve or keep performance between adaptive and non-adaptive stage. These results seem to suggest that online co-adaptation is an effective way to gain BCI control.

Keywords: electroencephalogram, brain, brain-computer interface, serious gaming, event related desynchronization.

Resumo

O cérebro é o principal órgão do sistema nervoso central. É composto por milhões de células altamente especializadas – os neurónios – que comunicam entre si e com outras células através de propagação de impulsos elétricos. Esses impulsos elétricos são responsáveis pela sensação, pensamento e movimento. Diferentes áreas do cérebro são responsáveis por diferentes processos. O córtex motor é parcialmente responsável pelo processamento da sensação háptica e do movimento voluntário. Para que haja movimento motor voluntário, várias áreas do cérebro comunicam entre si e com o córtex motor para planear, executar e atualizar o plano do movimento. Toda esta informação é transmitida por vias neuronais que ligam o cérebro aos músculos. Estas vias neuronais estão protegidas pela coluna vertebral, o motivo pelo qual lesões graves na coluna podem resultar em paralisia muscular.

Há diversos métodos de captar e analisar atividade cerebral. Estes métodos podem utilizar as diferenças de atividade metabólica em determinadas regiões do cérebro ou as diferenças de potencial causadas pelos sinais eletromagnéticos emitidos pelos neurónios. Estes métodos podem ainda ser invasivos - geralmente com maior precisão espacial e temporal - que requerem cirurgia, ou não-invasivos que embora tenham sinais com menor qualidade, são facilmente adquiridos, sem cirurgia. Neste projeto sinais de eletroencefalografia (EEG) foram utilizados, este método de aquisição foi o escolhido pela sua mobilidade, preço reduzido e facilidade com que se adquirem os dados, que fazem deste método o mais provável de ser posteriormente usado por utilizadores fora dos laboratórios.

Brain-Computer Interfaces (BCI, equivalente em Português a Interface Cérebro-Computador) recebem e interpretam sinais cerebrais e transformam-nos em comandos. Assim sendo, BCIs permitem aos seus utilizadores interagirem com o mundo que os rodeia utilizando apenas o pensamento. BCIs passivos utilizam o estado mental do utilizador (como ansiedade ou calma) para determinar uma ação. BCIs ativos são sistemas em que o utilizador utiliza processos mentais voluntários para controlar o software. As três categorias mais comuns de BCIs são *Event Related Potentials* (ERPs), *Steady-state Visually Evoked Potentials* (SSVEPs) e por fim, *Event Related Desynchronizations (ERDs)*. Entre estas categorias, BCIs baseados em ERDs são os únicos que podem ser usados espontaneamente, sem que o utilizador receba estímulos externos. Estes BCIs usam alterações espontâneas nos padrões de EEG provocadas por determinados processos mentais que causam sincronizações de dessincronizações neuronais. Um dos processos mentais mais frequentemente utilizado é *Motor Imagery* (MI), na qual o utilizador imagina o movimento de determinada parte do corpo. Diferentes estratégias mentais causam sincronizações e dessincronizações de sinapses em diferentes áreas do córtex e em diferentes frequências. A dessincronização neuronal em determinadas áreas do córtex faz com que a potencia de determinadas frequências diminua perto dessa mesma área. A deteção de diminuições e aumentos de potenciais nas frequências mu (10-12Hz) e beta (18-32Hz) estão na base de praticamente todos os sistemas de BCI controlados por imaginação motora.

Cada utilizador tem padrões de EEG específicos o que significa que os sistemas requerem calibração do BCI e treino do utilizador. Nos últimos anos sistemas de treino co-adaptativos têm mostrado resultados promissores. Neste sistema, o utilizador executa processos mentais e recebe feedback do sistema, que permite o utilizador aprender e adaptar-se ao sistema. Simultaneamente o sistema analisa os sinais

de EEG do utilizador e adapta o seu modelo ao utilizador. Investigação feita com estes sistemas tem resultados promissores, mas não há investigação que siga estes sistemas após o término da adaptação do sistema, ou seja, como num cenário real em que a intenção do utilizador não é conhecida pelo sistema. Isto leva à questão que este estudo tentar responder: será que um sistema co-adaptativo continuar eficaz após o término da adaptação?

Para testar esta hipótese foi desenhado um estudo no qual participaram 20 voluntários não remunerados. Para o sistema de BCI foram adquiridos sinais de 15 elétrodos situados sobre a área do córtex motor. As aquisições foram feitas num laboratório espaçoso e iluminado, com o voluntário sentado a aproximadamente 90 centímetros do ecrã.

Este estudo pretende aferir a eficácia de um sistema de BCI co-adaptativo após o término da adaptação. Para o efeito, foi desenhado um paradigma com feedback instantâneo com duração de aproximadamente 50 minutos. O sistema foi apresentado sobre a forma de um jogo. O jogo é uma variante do Whack-A-Mole, no qual os utilizadores devem martelar um fantasma que aparece e desaparece alternadamente do ecrã. Os utilizadores foram instruídos a executar MI quando o fantasma se encontrava no ecrã e a relaxar quando o fantasma desaparece.

O paradigma consiste em 5 rondas com 40 fantasmas (aproximadamente 10 minutos por ronda) intercaladas com pausas de 2-5 minutos. Nas primeiras 3 rondas, o sistema BCI adapta-se ao utilizador. O BCI atualiza o seu modelo a cada 5 fantasmas. O modelo utiliza os dados dos últimos 10 minutos, de forma ao modelo estar em conformidade com as técnicas mais recentemente usadas pelo utilizador.

O jogo tem alguns elementos visuais como um fantasma, um martelo e a energia e pontuação do jogador. O jogador ganha energia por descansar quando o fantasma está ausente. Toda a energia é perdida quando o utilizador martela sem o fantasma estar presente. A energia é convertida e pontos sempre que o utilizador martela o fantasma. O objetivo do jogador é acabar com a maior pontuação possível.

O paradigma foi feito sobre a forma de jogo em conformidade com investigação de psicologia educativa, que prova que sistema de treino mais entusiasmantes e motivantes têm melhores resultados do que sistemas mais simples e monótonos. O jogo tem feedback instantâneo que permite ao jogador aprender que estratégias são mais eficazes, tem um objetivo simples e claro, a experiência é feita num laboratório espaçoso e a explicação dos objetivos da experiência é feita de forma descontraída e informal. Tudo isto cria um ambiente familiar e confortável que promove uma melhor aprendizagem.

O processamento de sinal consiste primeiramente na extração da potência de determinadas frequências de sinal de EEG de determinadas áreas do cérebro. Dos elétrodos C3, Cz e C4 foram extraídas potências de sinais da gama alfa (10-13Hz), beta(16-24) e gamma (24-32Hz). Essas potências foram inseridas num classificador linear que criam um híper-plano que separa potências típicas de MI de potências típicas de não-MI. Esse classificador produz uma só variável que determina a probabilidade de o utilizador estar a executar MI. Quando esta probabilidade se mantém acima de 0.55 por 0.3 segundos o martelo levanta, e se se mantiver acima de 0.55 por mais 1.7 segundos o martelo martela.

Com base na forte correlação entre a exatidão do classificador e a alta pontuação dos jogadores pode-se aferir que o sistema de feedback funcionou como esperado. O nível máximo de performance alcançável sem controlo foi aferido por testes de permutação. Nestes testes, as potências utilizadas pelo classificador foram baralhadas de forma a não haver um padrão intencional. Nestas condições, a classificação terá uma exatidão inferior a 53%. De acordo com este raciocínio, pode-se concluir que 17 dos 20 utilizadores conseguiram ganhar controlo do sistema.

Quatro participantes voluntariaram-se para repetir a experiência múltiplas vezes. Três dos quatro participantes melhoraram a sua performance entre sessões. Estes resultados vão de acordo com o que seria de esperar tendo em conta investigação com utilizador de BCI a longo prazo. No entanto os resultados não são conclusivos dado ao baixo número de participantes.

Dos 18 utilizadores que conseguiram controlar o sistema, 17 mantiveram ou melhoraram a sua performance entre a fase adaptativa e não adaptativa. Estes resultados indicam que sistemas de BCI treinados de forma co-adaptativa são viáveis mesmo após o fim da adaptação. Estes resultados são especialmente entusiasmantes pois sistemas de classificação utilizados em cenários reais não têm informação prévia sobre a intenção do utilizador e, por consequência, não se podem adaptar. Criando um sistema eficaz após término da adaptação, esse sistema pode ser transposto para software de acessibilidade ou hardware como cadeiras de rodas que poderão ser controladas por pensamento. .

Palavras-chave: eletroencefalograma, cérebro, interface cérebro-computador, jogabilidade adaptada, dessincronização relacionada com eventos.

Contents

Acknowledgements **iii**

Abstract **v**

Resumo **vii**

List of Figures **xiii**

List of Tables **xv**

List of Abbreviations **xvii**

1 Introduction **1**
- 1.1 Brain and Movement . . . 1
  - 1.1.1 Brain signal acquisition . . . 1
- 1.2 Brain Computer Interface . . . 2
  - 1.2.1 Event Related Desynchronization based BCI . . . 3
  - 1.2.2 Co-adaptive BCI . . . 4
  - 1.2.3 Learn by playing . . . 5
  - 1.2.4 Overview of this Thesis . . . 5

2 Methods **7**
- 2.1 Study Participants . . . 7
- 2.2 Data Recording . . . 7
- 2.3 Experimental Design . . . 8
  - 2.3.1 Paradigm . . . 9
  - 2.3.2 Gameplay and Feedback . . . 11
  - 2.3.3 Play and Learn . . . 11
- 2.4 Signal Processing . . . 14
  - 2.4.1 BCI parameters extraction and online classification . . . 14
  - 2.4.2 Online co-adaptation . . . 15
  - 2.4.3 Artifact rejection . . . 15

3 Results **17**
- 3.1 Overall system in naïve subjects . . . 17

3.2 Adaptive and non-adaptive BCI 23
3.3 Evolution between sessions 25
3.4 Random Performance 27

4 Discussion **29**
4.1 Results analysis 29
4.1.1 Overall system 29
4.1.2 Adaptive vs non-adaptive system 30
4.1.3 Subject performance between sessions 30
4.2 Future Work 31
4.2.1 Error Related Potentials detection 31
4.2.2 The FORCe 31
4.2.3 Multiplayer improved gameplay 32

5 Conclusion **33**

Bibliography **35**

A Appendix - Study Information Sheet **39**

B Appendix - User Experience Form **43**

List of Figures

1.1 Average EEG responses to targets and distractors 3
1.2 ERD time-frequency maps for different kinds of motor imagery. 4

2.1 Electrode placement for data acquisition. 8
2.2 Single trial from the experimental paradigm. 9
2.3 Paradigm illustration. 10
2.4 Comparison between technical and game-like paradigm. 13
2.5 Overview of the pipeline used to process the data from raw EEG to *p_MI*. 14

3.1 Prediction curves and hammer hits histogram for the best, median and worst performers. 19
3.2 ERD maps for the best, median and worst performer. 21
3.3 Distribution of stars acquired per round (A) and true positive strike rate (B) relative to online BCI accuracy. 22
3.4 Distribution of self-reported control (A) and satisfaction with the system (B) relative to online BCI accuracy. 22
3.5 Comparison of different stages performance for subjects EF6, EG4 and EF3. 24
3.6 Histogram (A) of the differences in performance from adaptive to non-adaptive stage and scatter plot (B) of the same difference as a function of user online accuracy. 25
3.7 Subjects' Online Accuracy between sessions. 26
3.8 Histogram with the Online Accuracy from 10000 random performers. 27

4.1 Grand average EEG while charging the hammer intentionally and unintentionally. 32

List of Tables

1.1 Summary of neuroimaging methods. 2

3.1 Whack-A-Mole overall performance for 20 naïve subjects. 17
3.2 Difference in performance between adaptive and non-adaptive stage of the paradigm. . . 23
3.3 Difference in performance between sessions for subjects that volunteered multiple times. 25

List of Abbreviations

BCI	**B**rain **C**omputer **I**nterface
ECoG	**E**lectro**co**rtico**g**ragphy
EEG	**E**lectro**e**ncephalo**g**raphy
ERD	**E**vent **R**elated **D**esynchronization
ERP	**E**vent **R**elated **P**otential
ErrP	**Err**or **R**elated **P**otential
ERS	**E**vent **R**elated **S**ynchronization
fMRI	**f**unctional **M**agnetic **R**esonance **I**maging
ICA	**I**ndependent **C**omponent **A**nalysis
LFP	**L**ocal **F**ield **P**otentials
MEG	**M**agneto**e**ncephalo**g**raphy
MI	**M**otor **I**magery
MUA	**M**ultiple **U**nit **A**ctivity
NIRS	**N**ear **I**nfra**r**ed **S**pectroscopy
sLDA	**s**hrinkage regularized **L**inear **D**iscriminant **A**nalysis
SUA	**S**ingle **U**nit **A**ctivity
VEP	**V**isual **E**voked **P**otentials

1 Introduction

1.1 Brain and Movement

The brain is the main organ of the human nervous system. It is composed of billions of specialized cells - the neurons - that communicate among themselves and with other cells through neural electrochemical signals. These signals are responsible for sensations, thoughts and actions [1]. To perform any kind of voluntary action, a chain of electrochemical reactions between multiple neurons occurs. These reactions create a signal that is transmitted from the brain to an effector organ.

The gray matter on the outer surface of the cerebrum is the cerebral cortex. The cerebral cortex is divided in multiple areas like the visual, auditory, sensory somatic and motor cortex. The motor cortex (precentral gyrus) is located immediately anterior to the central sulcus. Action potentials initiated in this region control many voluntary movements, especially the fine motor movements of the hands [2]. Somatic sensory and motor cortex are side by side and the processing of haptic sensation and voluntary movement are strongly correlated which allow us to control and refine movements like grabbing or throwing so easily.

For voluntary motion to occur several areas of the brain need to work together with the motor cortex to plan, execute and update the movement [3] (e.g. how firmly should a muscle contract, how fast must the movement be, and how to react if something blocks the movement). All this information is sent via electrochemical signals in neural pathways that connect the brain and the muscles – the effector organ of voluntary movements. The neural pathways that connect the brain to most muscles are protected by the spinal cord. This is why spinal cord injuries often result in irreversible movement paralysis.

1.1.1 Brain signal acquisition

Brain activity can be measured and recorded using many different approaches. As examples, one can measure brain activity using differences in hemodynamics in the brain via functional magnetic resonance imaging (fMRI) or near-infrared spectroscopy (NIRS). In this work, we will focus on measuring brain activity using differences in electromagnetic activity in the brain.

Within electromagnetic brain signal acquisition methods, it is important to distinguish two categories: invasive and non-invasive. The methods in first category require a surgery to implant the electrodes beneath the cranium. Some highlightable invasive methods are single/multiple unit activity (SUA/MUA)

TABLE 1.1: Summary of neuroimaging methods [5]

Neuroimaging method	Activity measured	Temporal resolution	Spatial resolution	Risk	Portability
EEG	Electrical	∼0.05 s	∼10 mm	Non-invasive	Portable
MEG	Magnetic	∼0.05 s	∼5 mm	Non-invasive	Non-portable
ECoG	Electrical	0.003 s	∼1 mm	Invasive	Portable
Intracortical neuron recording	Electrical	∼0.003s	∼0.5 mm (LFP) ∼0.1 mm (MUA) ∼0.05 mm (SUA)	Invasive	Portable
fMRI	Metabolic	∼1 s	∼1 mm	Non-invasive	Non-portable
NIRS	Metabolic	∼1 s	∼5 mm	Non-invasive	Portable

which measure electric activity of single/multiple neurons respectively; local field potentials (LFP), which measures very local (from 600 μm to the centimeter scale) extracellular potential differences inside the brain [4]; and electrocorticography (ECoG) which records electrical activity from the surface of the cerebral cortex. The most common non-invasive methods are magnetoencephalography (MEG) and electroencephalography (EEG), which record noninvasively external magnetic fields and electric potentials, respectively. Table1.1 adapted from [5] compares some of the major aspects of the techniques mentioned above.

In this project, we used EEG signal. The main disadvantages of EEG are its low spatial accuracy and a poor signal-to-noise ratio compared to alternative modalities. On the other hand, it requires no surgery and the material required to measure EEG signal is comparatively affordable, light and mobile, making it the most suitable for real-life applications.

1.2 Brain Computer Interface

Brain-Computer Interface (BCI) technology translates brain signals into commands for the control of a device [6]–[9]. BCI users are thus enabled to interact with the environment by thought, or more generally speaking by mental processes. BCI can be used as a passive or active method of control. For a passive control, the software can use a person's state of mind to determine an action [10] (e.g. lowering/increasing a game difficulty if the user is stressed/bored). In an active control scenario, the user actively uses thoughts to control the software. The most popular paradigms for active BCI control use Event-Related Potentials (ERPs), Steady-state Visually Evoked Potentials (SSVEPs) or Event Related Desynchronizations (ERDs). In ERP paradigms the user typically stares at multiple sequential stimuli: some act as distractors and some as the target. When the subject is presented with a target stimulus among distractors, there is a positive peak around the Pz electrode (international 10–20 system) roughly 300 ms (p300) after the said stimulus. By identifying the peak, one can identify the user's target. This is called the oddball paradigm. This paradigm was used for the first time in 1988 to control a keyboard where columns and rows would flash sequentially. When the row or column that contains the desired character flashes, the user's brain response will be slightly different, and this is used to determine the

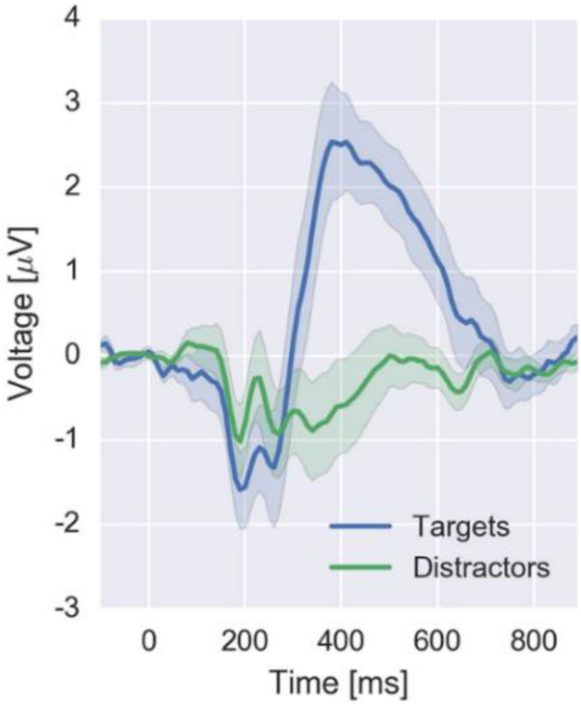


FIGURE 1.1: Time courses of the average EEG responses to targets and distractors at the midline electrode Pz. according to the international 10–20 system [12]

desired character [11]. Figure 1.1 shows the difference in EEG response to target stimuli and distracting stimuli. In this task the subject was presented with different colored stimuli and asked to count the number of times a specific target color appeared [12].

SSVEP protocols use visual stimuli modulated at different frequencies to generate brain potentials that fluctuate with the same frequency [13]. SSVEP stimuli are usually flashing images or checkerboard patterns. The frequency of the flashes can be detected in EEG of the visual cortex, so when the user is faced with different options flashing at different frequencies, the software can detect which option the user is focused on.

Both visual ERP and SSVEP are Visual Evoked Potentials (VEP) systems that require the user to gaze at flashing stimuli for the duration of the interaction. This may be extremely uncomfortable for long sessions and even dangerous for users with photosensitive epilepsy. The next section will be focused on a BCI alternative paradigm that requires no outside stimuli to the user.

1.2.1 Event Related Desynchronization based BCI

Event Related Desynchronization (ERD) based BCI uses the detection of changes in the spontaneous EEG signal. Different mental processes induce power decreases (ERD) or increases (event related synchronization, ERS) at different frequencies and in different areas of the brain. These differences can be measured and classified.

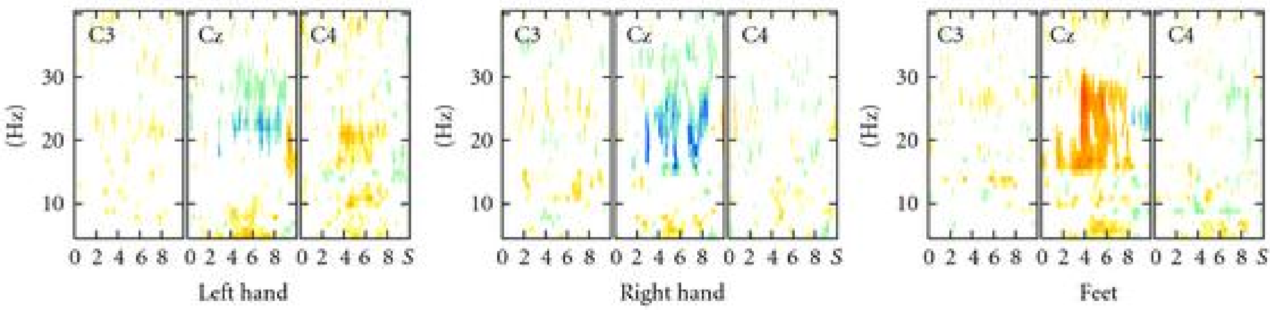


FIGURE 1.2: ERD time-frequency maps for different kinds of motor imagery. In this figure, red indicates higher levels or ERD, meaning lower power on the given frequency, time and electrode [15]

Motor imagery (MI), i.e. the kinesthetic imagination of movement, is a mental process commonly used to modulate EEG rhythms and induce characteristic activity patterns in the EEG signal [14]. For instance, Figure 1.2 shows a time-frequency map where different areas of the brain show ERD and ERS depending on the type of motor imagery. Usually, feet imagery causes midcentral focused μ(10-12 Hz) and β(18-32 Hz) ERD and bilateral beta ERS. Hand imagery has an antagonistic behavior, causing ERS in the midcentral area and ERD in the lateral area contralateral to the hand [15].

As a result, an MI BCI system allows its user to control a software by imagining movement, with no need for displaying external stimuli to the user. This means that MI has a great potential for self-paced BCI, where the user may pick the timing of his action.

MI typically has a lower accuracy than the VEP alternatives [10]. Fortunately, the ability to produce reliable EEG patterns with MI is trainable and users tend to increase accuracy over time [16]–[18]. If MI performance is trainable, one may wonder what factors influence the training effectiveness. Perhaps the BEST known factor to influence user performance is feedback. It has been consistently shown that feedback is correlated with an increase in performance [17], [19]–[21]. Showing users how well they performed, gives them information about what strategies work better on them and also a sense of motivation that keeps the users more engaged [21]. More details on how to calculate ERD are given in the *Signal Processing* section.

1.2.2 Co-adaptive BCI

EEG patterns are user-specific and change non-linearly over time. Consequently, BCIs do not work "out of the box" and need user-specific adaptation. Lately, online co-adaptive training approaches achieved promising results [21], [22]. Online co-adaptation means that user performs mental processes and receives online feedback on the detected EEG pattern. This allows the user to learn from the BCI (neurofeedback training). Simultaneously the BCI analyzes the ongoing EEG and adapts its model parameters to optimize pattern detection, which allows the BCI to learn from the user (machine learning). Pattern generation and pattern detection mutually adapt online. Faller et al. [21] reported 10 out of 12 naïve (first time) BCI users were able to make binary decisions with accuracy higher than 70% after 60 to 80 minutes of online co-adaptive training. Schwarz et al. [23] reported that 9 naïve users achieved an average accuracy of 81.4% within 60 minutes. Kobler [24] reported that 10 naïve users achieved average accuracy of 83.5% within 24 minutes. In another report by Faller et al. [19] 18 out of 22 severely impaired end-users attained BCI control levels that were better than random within 24 minutes (average accuracy of 74%). Moreover, Faller implemented an online sham feedback experiment to test whether disabled users would be able to interact via BCI after the training. Offline analysis suggests that half of the users would succeed. Research to date has focused on enhancing co-adaptation methods. To the extent our knowledge, no study evaluated the online BCI control proficiency of users that went through online co-adaptive training. The current study evaluates BCI performance during and after online co-adaptive training.

1.2.3 Learn by playing

The current study also addresses another issue of BCI. There is increasing awareness that current experimental BCI training paradigms are sub-optimal for users with respect to educational psychology and instructional design [17], [25]. Key-factors for successful human training include motivation, challenge and clarity of training goal. Conventional BCI training paradigms are boring, monotonous and graphically not appealing [26]. To promote training success, in this study, co-adaptation is embedded within a computer game. A variant of the Whack-A-Mole game was implemented. Whack-A-Mole is a game in which players use a club to hit toy moles back into their holes. The resulting training environment is intuitive to play, the goal is clear, and the design is pleasing.

1.2.4 Overview of this Thesis

This thesis reflects the work of a 6-month Master's project during an internship at the Institute of Neural Engineering (INE) of the Graz University of Technology, in Austria.

The project aims to test if a co-adaptive MI-based BCI maintains its' effectiveness after the adaptation stops. To test it, the BCI was integrated into a game with online feedback and played by 20 naive subjects.

Preliminary results were used to write a paper for the *2018 International Conference on Cyberworlds* titled "Are Online Co-Adaptive Sensorimotor Rhythm Brain-Computer Interface Training Paradigms Effective?" [27]. The most recent results are being used to write a new paper to be submitted in *Frontiers in Neuroscience*.

2 Methods

2.1 Study Participants

Twenty naïve volunteers (mean age 26 $\pm$ 3 years, 5 females) participated in this study. Volunteers were without any known medical condition, had normal or corrected to normal vision and entered the study voluntarily without monetary remuneration. At the beginning of the study, each participant was briefed about the aim of the study and presented with an information sheet attached in *Appendix A*. All gave written informed consent to participate. The study was conducted in accordance with the relevant guidelines for ethical research according to the Declaration of Helsinki.

2.2 Data Recording

EEG was recorded from 15 Ag/AgCl electrodes, (over locations FC3, FCC1h, FCC2h, FC4, C5, C3, C1, Cz, C2, C4, C6, CCP1h, CCP2h, CP3, and CP4) placed over the motor cortical area. These electrodes were integrated into a customized 64-channel cap (waveguard, ANT Neuro, Enschede, Netherlands). See Figure 2.1 for cap layout.

For further research, the EEG was recorded from 3 extra electrodes over location Fz, Pz, and Oz. These electrodes, together with Cz, where used in previous research [28] to observe Error Related Potentials (ErrPs). These potentials will be further discussed in the *Future Work* section.

All signals were referenced to electrode CPz and kept with impedance under 20kΩ. The ground electrode was placed at AFz. EEG was recorded using a biosignal amplifier (eegosports, ANT Neuro, Enschede, Netherlands) and sampled at a rate of 512 Hz.

All recordings were made in a well illuminated spacious room shared between the subject and 2 researchers. Subjects were sitting in a comfortable chair approximately 90 centimeters from the screen.

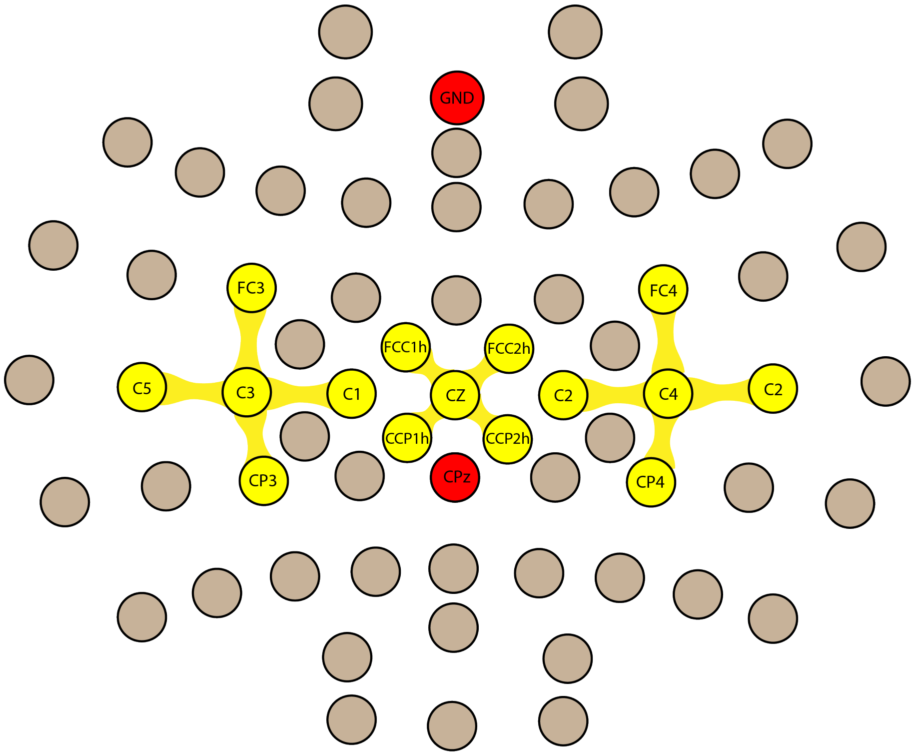


FIGURE 2.1: Electrode placement for data acquisition. Laplacian referencing is marked in yellow. Ground and Reference electrodes are marked in red.

2.3 Experimental Design

An experiment was developed to determine if a co-adaptive BCI system remains effective after the adaptation stops. In this experiment the subjects use a BCI system with online feedback for approximately 50 minutes. During the first 30 minutes, the BCI parameters adapt to the user, while in the last 20 minutes the BCI parameters are static and based on each subject's previously recorded data.

Before the data recording starts, the subjects are informed about the goals and methods of the experiment while the EEG cap is set up.

The BCI system is presented to the user in the form of a game. Participants were seated in front of a computer screen with the task to play a variant of the Whack-A-Mole game. Users were asked to perform hand motor imagery (MI) whenever a ghost character was shown on the screen. The game was implemented in the cross-platform game engine Unity (Unity Technologies, San Francisco, USA). Communication between game and BCI was implemented by the lab-streaming-layer (LSL) network protocol From the Swartz Center for Computational Neuroscience (SCCN) which ensures the unified collection of measurement time series (freely available in GitHub[1]).

[1]https://github.com/sccn/labstreaminglayer

All subjects were asked to fill up a user experience form twice: after 30 minutes of recording and at the end of the session. Each recording session lasted about 90 minutes including EEG montage, subject education, data recording and periodic pauses.

2.3.1 Paradigm

Participants were seated in front of a computer screen with the task to play a variant of the Whack-A-Mole game. Users were asked to perform hand MI whenever a cartoon ghost character was shown on the screen (Figure 2.2) and to relax when it disappears. The ghost would appear, play an audio cue and stay visible for 7 seconds. A random break between 7 to 8 seconds was presented before the next ghost appeared. In this document, we call a single trial to the time between the disappearance of two consecutive ghosts.

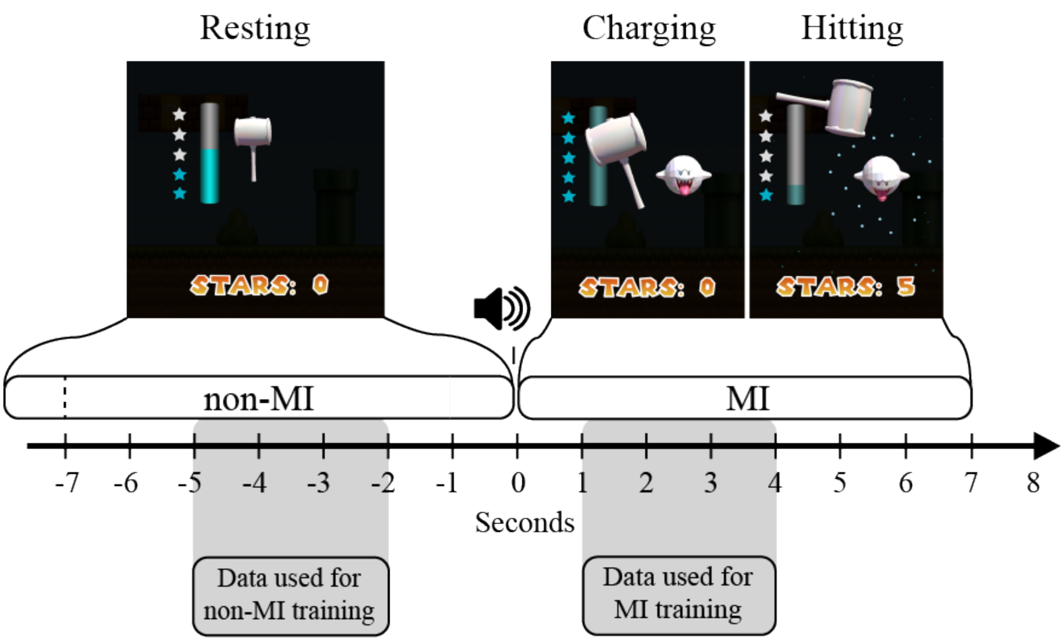


FIGURE 2.2: Single trial from the experimental paradigm.

Each session lasted 200 trials, divided into 5 rounds of 40 trials each. Between each round, subjects were free to take a break to move, ask questions, drink and eat. The duration of this pause was decided by each subject and would take approximately 1-5 minutes. In the first 3 rounds (120 trials) the BCI parameters adapted to the user. The online co-adaptive training is divided in calibration and recurrent adaptation stage [21]. The aim of the calibration stage is to collect EEG trials to compute the first set of BCI parameters. During calibration, sham feedback is provided to the user, i.e. the game is playing automatically with a predefined accuracy. When the first set of BCI parameters is available, training switches to the recurrent adaptation stage. Now the user is in control. The BCI analyzes the ongoing

EEG and recurrently recomputes BCI parameters to optimize detection. The following optimization policy was adopted:

1. *Calibration stage:* The first five trials in the first round were used to train the first classifier.
2. *Recurrent Adaptation stage:* every 5 trials a new classifier was trained. To give more recent features changes and higher impact, only the last 40 trials were included to retrain the system. Up to the fortieth trial (first round) all trials were included. The output of the classifier is a stream with a number between 0 and 1 that asserts the probability (*p_MI*) of the subject being performing MI at a given moment. More details on the data processing will be given in the *Signal Processing* section

In the fourth and fifth rounds the system is no longer adapting to the user. Instead, it uses the information from trial 80 to trial 120 (third round) to classify the subjects signal and give feedback in the same way as in the previous rounds. The classification of this stage resembles more closely what a real-life scenario would look like since the system has no pre-labeled information to train with. A scheme of the paradigm is shown in Figure 2.3.

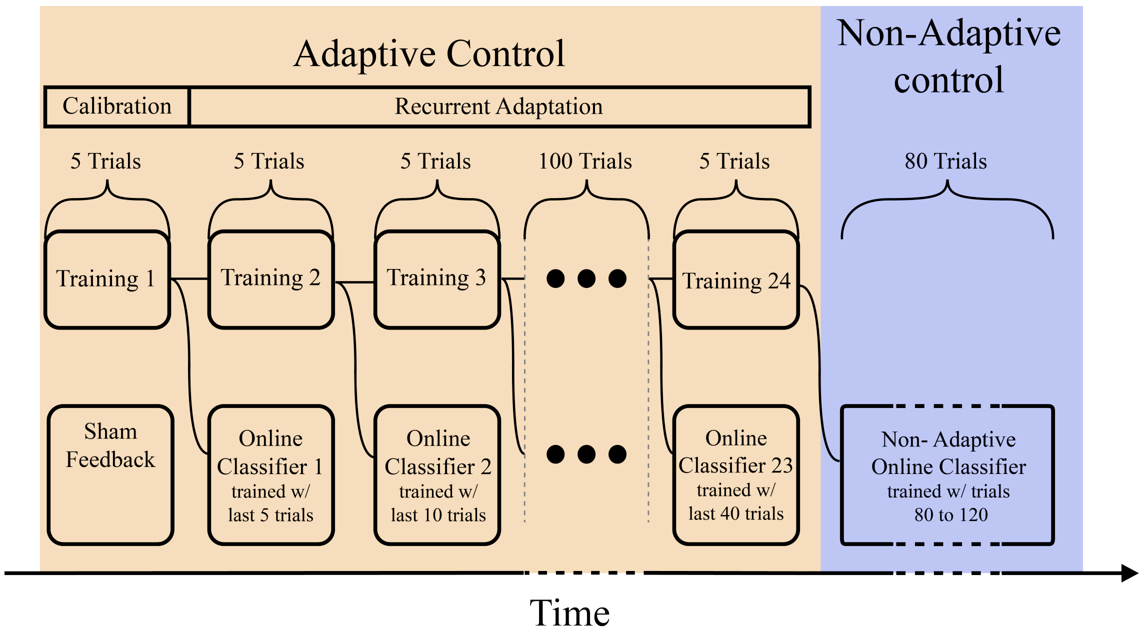


FIGURE 2.3: Paradigm illustration.

At the end of the experiment, subjects were asked to fill a form on user experience where they report how they felt physically and mentally during the experiment, how satisfied they were with the system and how much did they feel in control in both adaptive and non-adaptive stages of the game. This form is shown in *Appendix B*.

2.3.2 Gameplay and Feedback

The data is recorded while the subject plays a game, as mentioned in the previous section. In this game there are a few elements:

- **An intermittent cartoon ghost** that cues the user on when to perform MI.
- **A blue energy bar and stars** that indicates how much energy the subject got from resting.
- **An orange points counter and bar** that indicates how much stars did the user collect so far and how well he performed overall.
- **A hammer** that is controlled directly by the BCI and gives feedback to the user on his MI performance.

The hammer has three states: resting, charging and striking and the game mechanics goes as following:

When *p_MI* is under 0.55 the hammer is in the resting state. During this state, the user recovers energy which is displayed in the blue energy bar on the left. The subject reaches maximum energy by resting for 5 seconds. When the player is successfully resting, the energy bar becomes a lighter blue, increases and lights up a star for every 20% energy recovered (up to 5).

When *p_MI* is over 0.55 for more than 0.3 uninterrupted seconds, the hammer increases in size and goes up, indicating the user that MI is being recognized. During the charging state, no energy is lost nor recovered.

When *p_MI* is over 0.55 for 2 uninterrupted seconds the hammer strikes. If the ghost is in the screen, the energy is reset to 20% and the current stars are converted into points. This scenario will be called a hit. If the ghost is not on the screen, the energy and stars are reset to 0.

This set of rules encourages the subject to avoid performing MI when the ghost is not present (recovering energy) and to continuously perform MI when the ghost is present, even after the first hit, since a second and third hit on the same ghost will still provide points. The threshold for hammer charging was set to 0.55 instead of 0.5 to lower false positives that result in more on-screen movement and could disturb the users' concentration.

In the bottom of the game it is displayed the number of points (stars) collected along with a bar that indicates overall performance, which is non-linearly inflated so that subjects with low performance are still motivated. This and other decisions regarding user instructions and recording conditions were meant to improve the learning experience. This will be discussed in the next section.

2.3.3 Play and Learn

Reading the work of Lotte et al. [17] it becomes evident that many BCI training paradigms are sub-optimal regarding psychology and instructional design. To fight this tendency, we took measures to improve the motivation and learning process of the subjects.

First, each subject was instructed calmly in an informal and comfortable scenario. Each subject was given a lemon to squeeze and to feel its texture. Then they were asked to mentally reproduce the motor and haptic feeling of squeezing the lemon (not the visual image of doing it), in both hands, and instructed on other techniques of motor imagery. They were advised to pick a strategy that feels haptically vivid when imagined, and that they could keep consistent for the next hour, since the classifier works at its best with consistent strategies. After this, subjects were informed about EEG artifacts and how they affect the data. With an online EEG visualizer, subjects saw how their muscular contractions and eye blinking effect the signal. Subjects were asked to avoid blinking with a pattern (e.g. blinking every time the ghost appears) but to blink as they needed. All this was transmitted to the user, so that the procedure is clear and understandable.

Second, the paradigm was made into a simple yet stimulating game. The cartoon ghost resembles a Super Mario character (Boo), and the recording was made in a shared spacious room, instead of in a noise-isolation chamber. All of this made the environment comfortable and fun, hence ideal for learning.

Finally, the online feedback (hammer movement) was discrete and clear, enabling the user to learn over time, and the performance feedback (points counter and bar) were non-punishing yet gave a sense of purpose to the subjects.

To make clear the importance of 1) making the instructions understandable and simple 2) setting up fun goals in a familiar environment and 3) giving online feedback to the user, I challenge the reader to mentally perform Paradigm 1 and 2 in Figure 2.4.

Using a paradigm that resembles a game more than a test partially justifies the high levels of satisfaction reported in the users' experience form. This kind of paradigm also led to some users being eager to repeat the experiment and improve their performance.

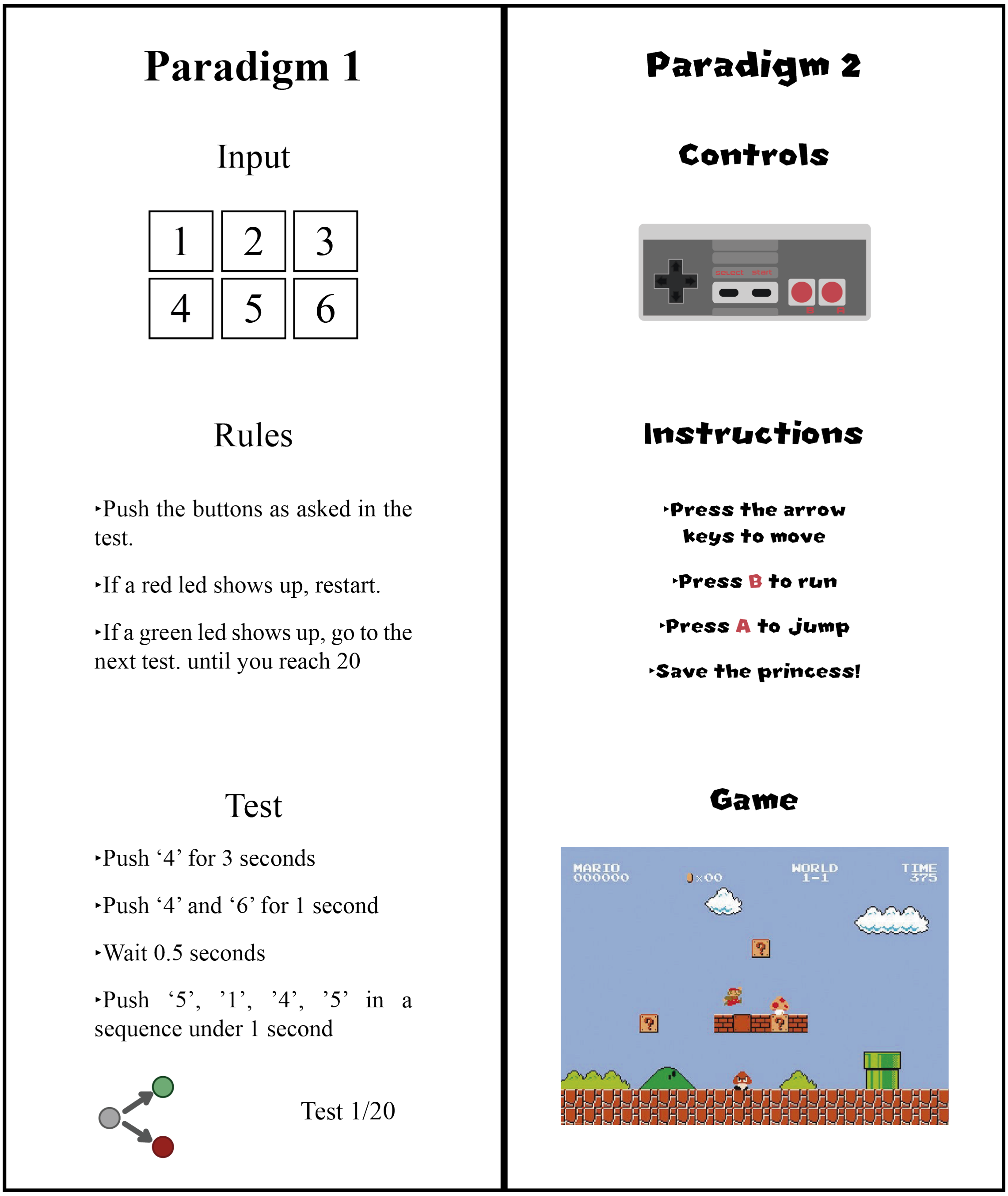


FIGURE 2.4: Two paradigms with similar control and task. Paradigm 1 (left) is technical and repetitive. Paradigm 2 (right) was voluntarily played by millions of people since 1985.

2.4 Signal Processing

2.4.1 BCI parameters extraction and online classification

Logarithmic band power features of α (10-13 Hz), low β (16-24 Hz) and high β / low γ (24-36 Hz) activity were extracted from three Laplacian derivations around C3, Cz and C4. Laplacian derivations are the signal from the central electrode subtracted by the average signal of the surrounding electrodes (Figure 2.1). Band power was computed by squaring samples of the filtered signal (fifth order Butterworth band-pass filters on α, β and γ frequency ranges), averaging the squared samples of the past second and computing the logarithm of the averaged value. This results in a feature vector of 9 features (three frequency ranges for each channel) that describe the last second of data. A shrinkage regularized linear discriminant analysis classifier (sLDA) was used to detect MI. LDA is a linear hyperplane classifier that divides the feature space in two parts. The separating hyperplane is positioned based on the co-variance of class-specific features. This method minimizes features variance within a class and maximizes the average difference in features between different classes. Shrinkage allows achieving well-conditioned co-variance matrices when data is high-dimensional, and few training data is available [29]. The sLDA output was the probability *p_MI* that the current feature vector belongs to the MI class. This process is repeated eight times per second, which is the update rate of the feedback, meaning that the feedback on the last second of data is updated every 125 ms. As a reference, the average reaction time between a visual stimulus and a button press is about 240 ms [30]. A pipeline/scheme of the parameters extraction and classification is shown in Figure 2.5

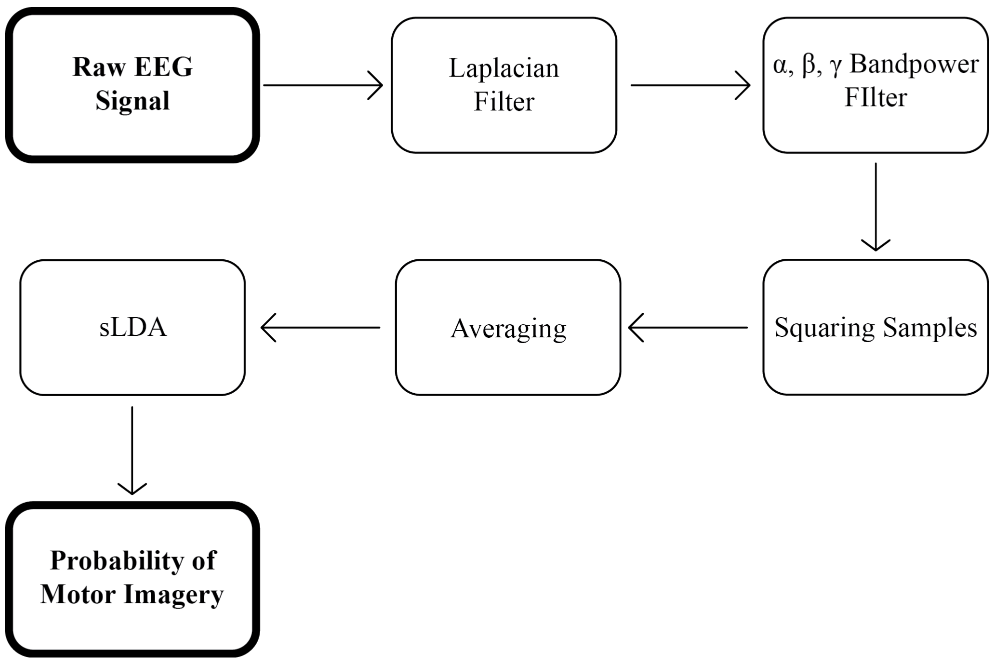


FIGURE 2.5: Overview of the pipeline used to process the data from raw EEG to *p_MI*.

2.4.2 Online co-adaptation

As mentioned in the previous section, online co-adaptive training is divided in calibration and recurrent adaptation stage. During the calibration stage, EEG is collected and the BCI computes the first set of parameters and calculates the first sLDA hyper-plane. When the first set of BCI parameters is available, training switches to the recurrent adaptation stage. The BCI analyzes the ongoing EEG and periodically recomputes BCI parameters to optimize detection. These two stages go as following:

1. *Calibration stage:* The first five trials in the first round were used to train the first sLDA. Six 9-features vectors were extracted from each trial. Feature vectors extracted at times t = 2, t = 3 and t = 4 seconds relative to the appearance of the ghost were used to characterize MI. Feature vectors extracted at t = -4, t = -3 and t = -2 seconds characterized non-MI. Each feature is extracted over one second, meaning that the time used to train the MI is the interval between 1 and 4 seconds relative to the ghost's appearance and to train the non-MI the interval between -5 and -2 seconds.

2. *Recurrent Adaptation stage:* Every five trials a new sLDA classifier was trained as described above. To give more recent feature changes a higher impact, only the last 40 trials were included to retrain the sLDA [16]. Up to the 40^{th} trial (first round) all trials were included. Preliminary test showed that using all acquired data would damage the classifier performance in the long run. This may be explained by the subjects' tendency to make small strategy changes over time. If the training window is too big, the sLDA cannot handle small changes ins strategy, if too small it becomes unstable and gives no time for the user to learn. 40 trials was a probationary window size picked from analyzing offline data, since picking an empirically optimal window would require extensive testing on different window sizes with multiple subjects.

2.4.3 Artifact rejection

Artifact rejection was made every time the sLDA was retrained. Eye movement and muscle contraction tend to increase power in α [31], and β and γ band [32] respectively. 40 trials had 120 features vectors with 9 features each for each condition (MI and non-MI). The 9 individual features were analyzed, and a vector would be rejected if it had one feature bigger than the mean plus three times the standard deviation of the same feature. This means that if in a given second one of the features (e.g. the beta band power in C3 Laplacian) was unusually high then all the features from that second were rejected.

3 Results

3.1 Overall system in naïve subjects

The performance of the overall system is summarized in Table 3.1. It displays, for each user, the score per round (number of collected stars), the percentage of strikes that hit (true positive (TP) hits), the accuracy of the classifier and self-reported level of control of the game during the whole session.

TABLE 3.1: Whack-A-Mole overall performance for 20 naïve subjects.

Subject Code	Online Accuracy (%)	Trial Accuracy (%)	TP Strikes(%)	Stars per Round	Control*	Satisfaction**
ED6	87.0	98.2	89.5	222	10	9
EE5	81.8	93.9	87.0	203	8.5	9
EG2	76.6	89.8	84.7	182	7	9
EG8	76.6	90.3	84.5	179	7.5	8
EF9	74.7	85.9	81.5	197	7	9.5
EF3	73.2	87.0	78.4	171	8.5	9
EG3	71.0	80.6	75.6	161	8	8.5
EG4	67.2	80.1	75.4	163	7.5	10
EF6	66.9	73.9	68.2	156	8.5	6.5
EE3	63.6	72.6	65.0	109	6.5	6.5
EF1	62.9	69.6	70.6	99	7	7.5
EF7	62.4	69.5	70.1	124	8	10
EF8	61.6	70.6	72.5	108	5.5	5
EF5	61.3	68.8	68.6	124	4.5	5
EG7	58.1	61.4	62.8	67	4.5	6
EF4	56.9	61.6	61.7	92	4.5	5.5
EG1	54.6	55.0	56.8	94	2.5	7
EG5	53.9	55.0	60.1	52	6.5	7
EG6***	53.0	55.5	57.6	50	3	4
EF2***	51.9	50.9	50.1	63	6.5	8
Mean	66.4	74.7	72.1	132.4	6.9	7.6
Std	9.3	13.3	10.5	48.3	1.9	1.6

*Question: "Did you feel in control of the game?"
Answer: 1/Very little ... 10/Very much
**Question: "Overall satisfaction with the system"
Answer: 1/Unsatisfied ... 10/Satisfied
*** Control proficiency not significantly better than chance (α = 0.01).

The accuracy of the system is calculated using the following formula

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{3.1}$$

Where TP, TN, FP and FN stand for True Positive, True Negative, False Positive and False Negative respectively. Accuracy was chosen as a measure of performance over Cohen kappa coefficient [33] as it is more intuitive to interpret, and it carries the same information given that the number of true-states and false-states are the same. To ensure this, this accuracy refers to the classifications between 1 and 7 seconds after the appearance and disappearance of the ghost. Two accuracy values are presented in Table 3.1, the first (Online Accuracy) takes the average classification over each second and labels it as true or false (one classification per second). This accuracy is the one that most accurately describes the game's input from the BCI. The second (Trial accuracy) does the same, averaging classifications between seconds 1 and 7 after the appearance and disappearance of the ghost (one classification per trial).

Perfect control results in an online accuracy, trial accuracy and TP strikes equal to 100%. This means no false positive strikes occurred and each ghost was hit three times. This results in a maximum score of 280 stars per round. Random performance level was assessed by permutation tests and resulted in an Online Accuracy = 53.4% and Stars per round = 64 (α = 0.01). A static hammer results in a score of 0 stars. A hammer that is constantly striking achieves a score of 80 stars per round, but this scenario is impossible as, by definition, during the adaptation, the sLDA hyperplane is placed between each class average feature vector. Consequently the sLDA cannot constantly output a prediction *p_MI* over 0.50 because the sLDA hyperplane constantly adapts(and the game threshold is 0.55).

A closer look on the higher, median, and worst performer (according to online BCI accuracy) shows the difference both directly in the classifier output and the game output for difference performances tiers. Average sLDA probability *p_MI* curves, and histograms with boxplots of hammer strike both as a function of time for users with different online accuracy are shown in Figure 3.1.

The subjects explored in Figure 3.1 have visible differences in BCI classification over time. The BCI detects and classifies ERD. Figure 3.2 shows the ERD in the 3 Laplacians for the 3 subjects for frequencies up to 32 Hz from 4 seconds before to 8 seconds after the appearance of the ghost.

The subjects do not get direct feedback from the BCI, instead, they assert their performance trough in-game hammer movement and final score. The relationship between score per round and BCI accuracy is shown in Figure 3.3 in a scatter plot with every subject's accuracy and score per round.

During the third break and after the recording the subjects answered a form where they could report their satisfaction a degree of control in the game. On a scale of 1 to 10 the average self-reported control and satisfaction was 6.9 and 7.6. There is a correlation (p-value <0.001) between user online accuracy and self-reported control and satisfaction. The relation between users' reports and performance can be visualized in Figure 3.4.

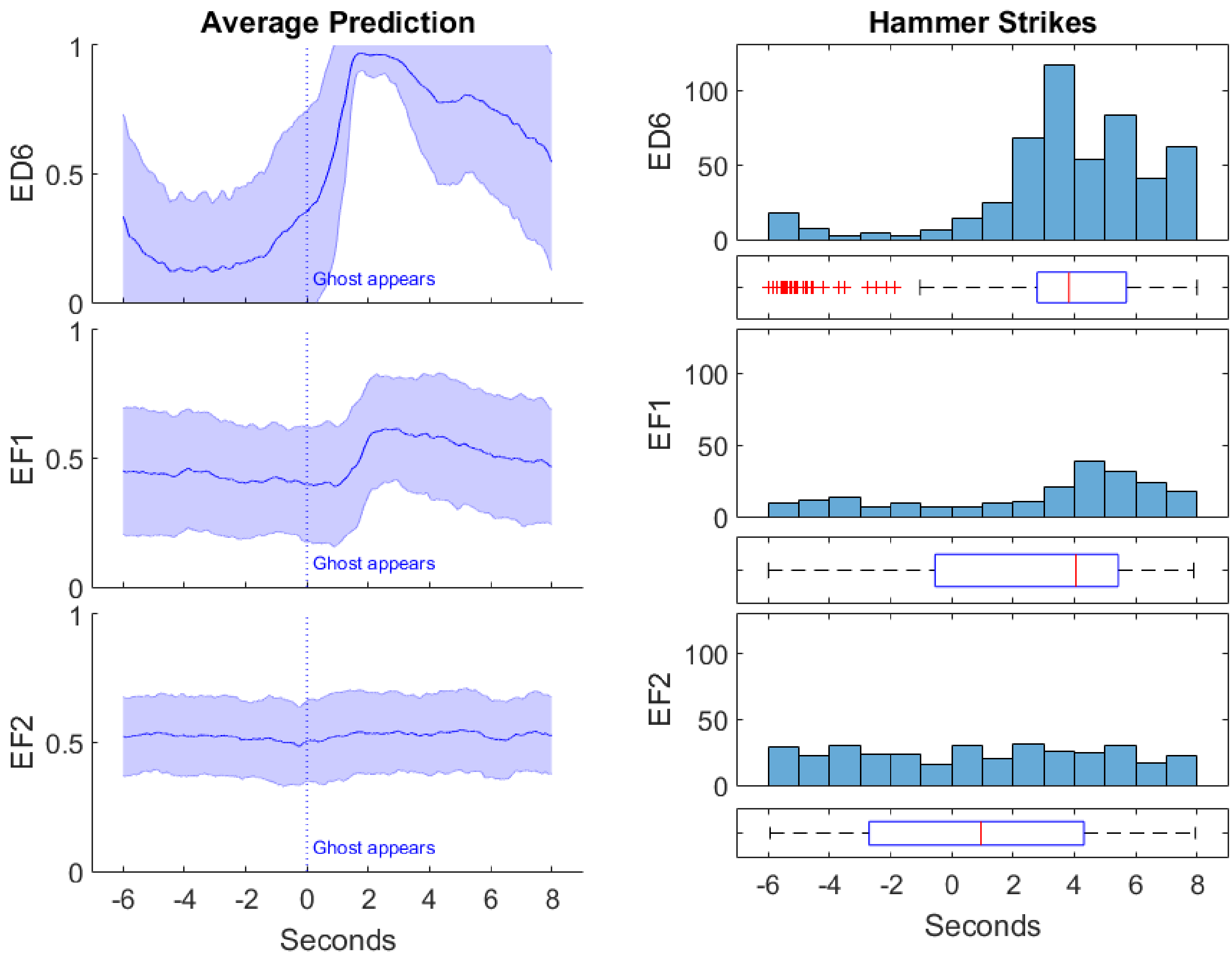
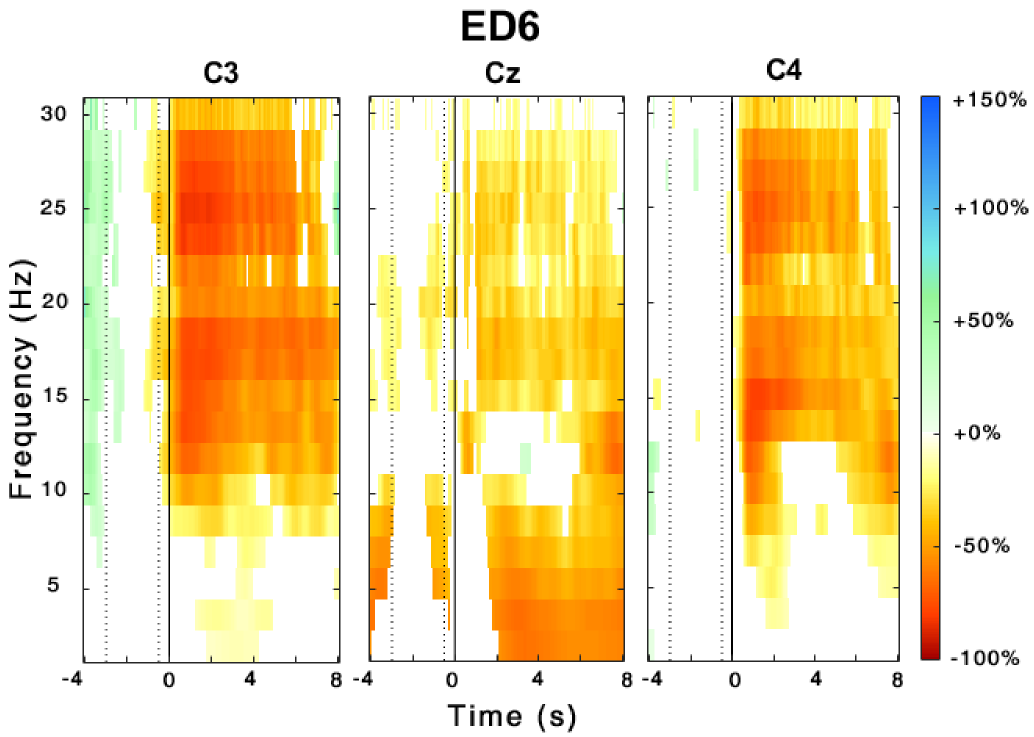
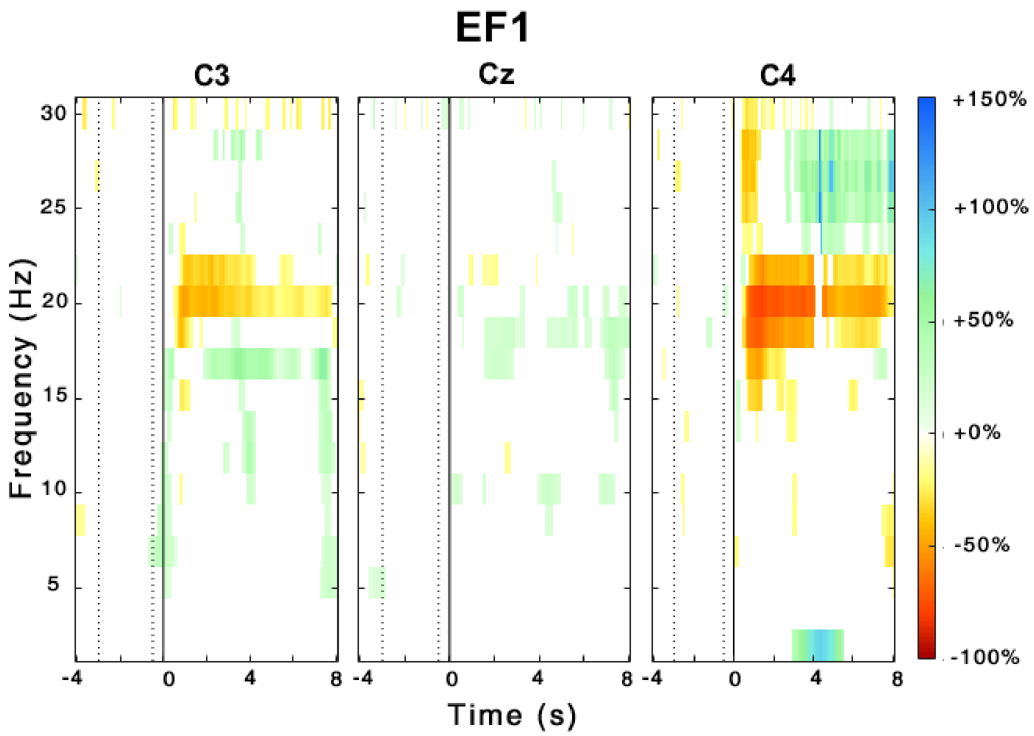


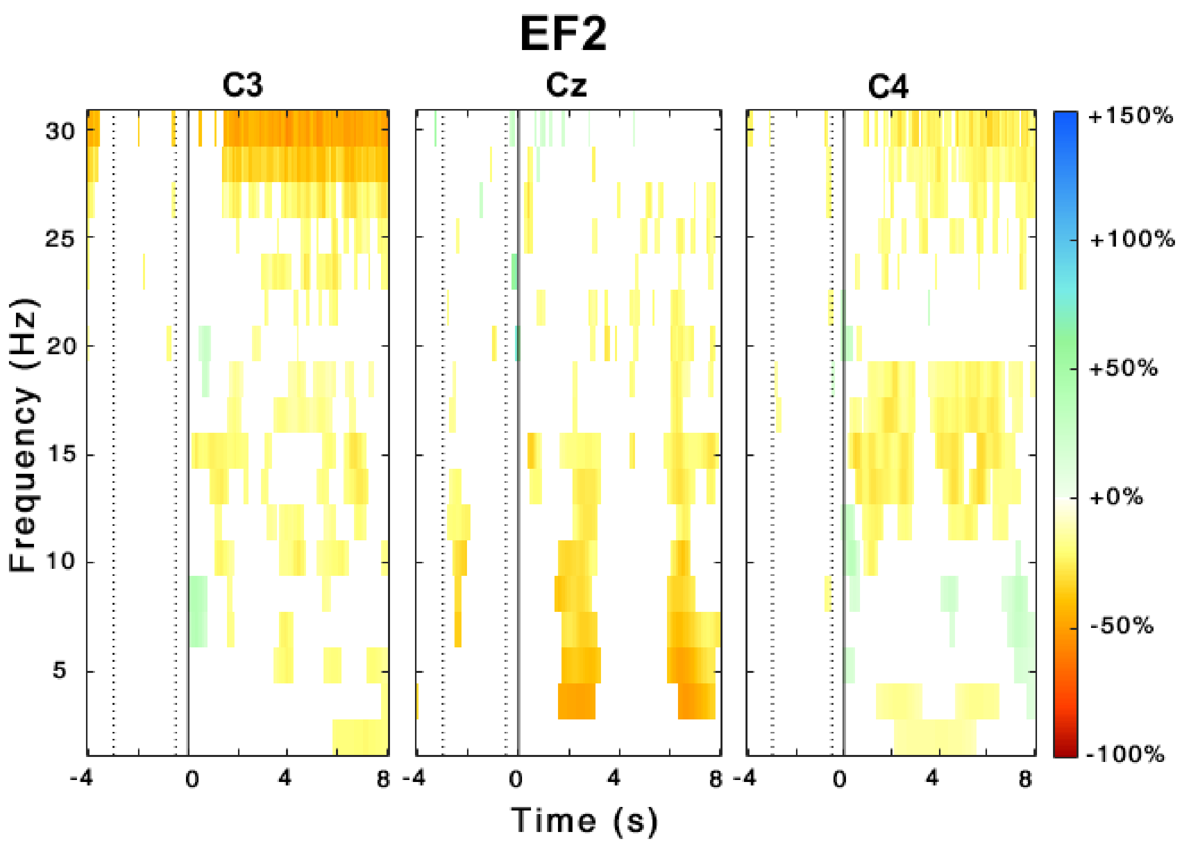
FIGURE 3.1: Results for the best (top), median (center) and worst (bottom) performers. Average online prediction (left). Histogram and boxplot of hammer strikes (right). The curves on the left show the average ± standard deviation (highlighted area) sLDA prediction *p_MI* curves. The graphs on the right show hammer strikes. The timing on the x-axes is relative to the appearance of the ghost (second 0).


ED6
C3
Cz
C4
Frequency (Hz)
30
25
20
15
10
5
-4
0
4
8
Time (s)
+150%
+100%
+50%
+0%
-50%
-100%

(A)


EF1
C3
Cz
C4
Frequency (Hz)
30
25
20
15
10
5
-4
0
4
8
Time (s)
+150%
+100%
+50%
+0%
-50%
-100%

(B)



(C)

FIGURE 3.2: ERD maps for the best (A), median (B) and worst (C) performer. These maps use the average of the bandpower in the interval between -3 and -0.5 seconds (dotted lines) as baselines. The color scale represents an increase (blue) or decrease (red) in power for each frequency relative to the baselines.

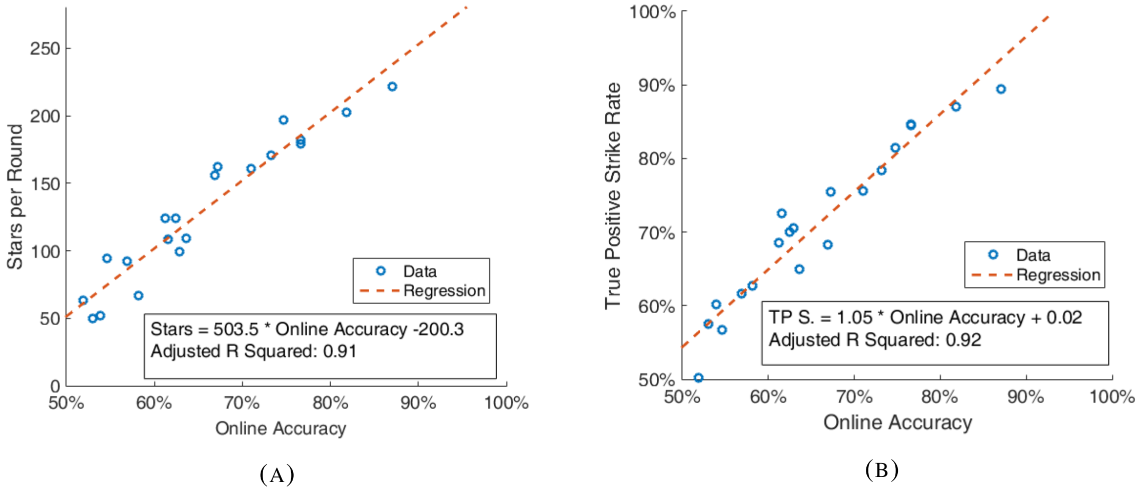


FIGURE 3.3: Distribution of stars acquired per round (A) and true positive strike rate (B) relative to online BCI accuracy.

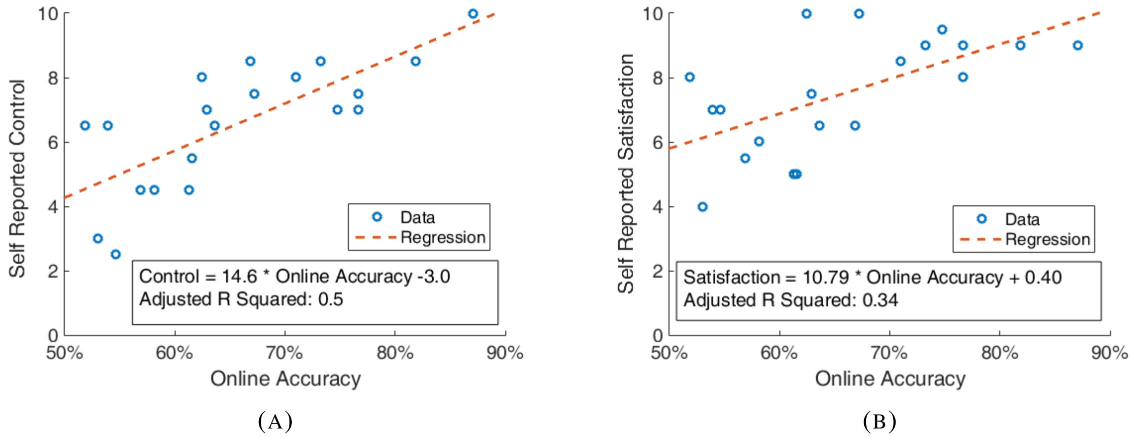


FIGURE 3.4: Distribution of self-reported control (A) and satisfaction with the system (B) relative to online BCI accuracy.

3.2 Adaptive and non-adaptive BCI

The BCI had an adaptive stage (round 1 to 3 in-game) and a non-adaptive stage (round 4 and 5 in-game). The difference in BCI accuracy, Stars collected and self-reported control and satisfaction between both stages are summarized in Table 3.2.

TABLE 3.2: Difference in performance between adaptive and non-adaptive stage of the paradigm.

Subject code	Online Accuracy (%)	△ Online Accuracy (%)	△Trial Accuracy(%)	△ Stars per Round	△ Control
ED6	87.0	-2.0	2.2	8	0
EE5	81.8	5.9	3.9	19	1
EG2	76.6	0.7	-3.5	3	2
EG8	76.6	2.0	2.4	-19	-1
EF9	74.7	-2.7	-7.1	-8	-2
EF3	73.2	-8.8	-3.4	-28	-1
EG3	71.0	5.0	0.9	36	2
EG4	67.2	0.0	3.3	35	1
EF6	66.9	10.7	13.8	53	1
EE3	63.6	-1.7	1.3	-1	-1
EF1	62.9	-3.0	-1.5	24	0
EF7	62.4	6.1	4.5	32	0
EF8	61.6	5.0	9.3	45	3
EF5	61.3	7.2	16.7	66	1
EG7	58.1	2.3	4.2	-27	3
EF4	56.9	1.0	2.9	19	3
EG1	54.6	-2.3	-4.4	-4	-3
EG5	53.9	-0.9	2.3	-11	1
EG6	53.0	1.1	8.5	16	0
EF2	51.9	-1.0	-4.1	-3	3
Mean	66.4	0.7	1.9	12.2	0.4
Std	9.3	4.1	5.7	24.1	1.8

Among all subjects, a few have notably high performance differences between stages. Subject EF5 and EF3 have the biggest positive and negative differences. Average probability *p_MI* curves, and histograms with boxplots of hammer strikes for these subjects in both stages are shown in Figure 3.5. Subject EG4 is also shown in the figure as a reference since it is the subject with the least difference between stages.

The difference in performance between the first and second stage are shown in a histogram in Figure 3.6. These differences are also plotted as a function of user performance. Statistical analysis shows no significant relation between overall user performance and an increase of performance between stages.

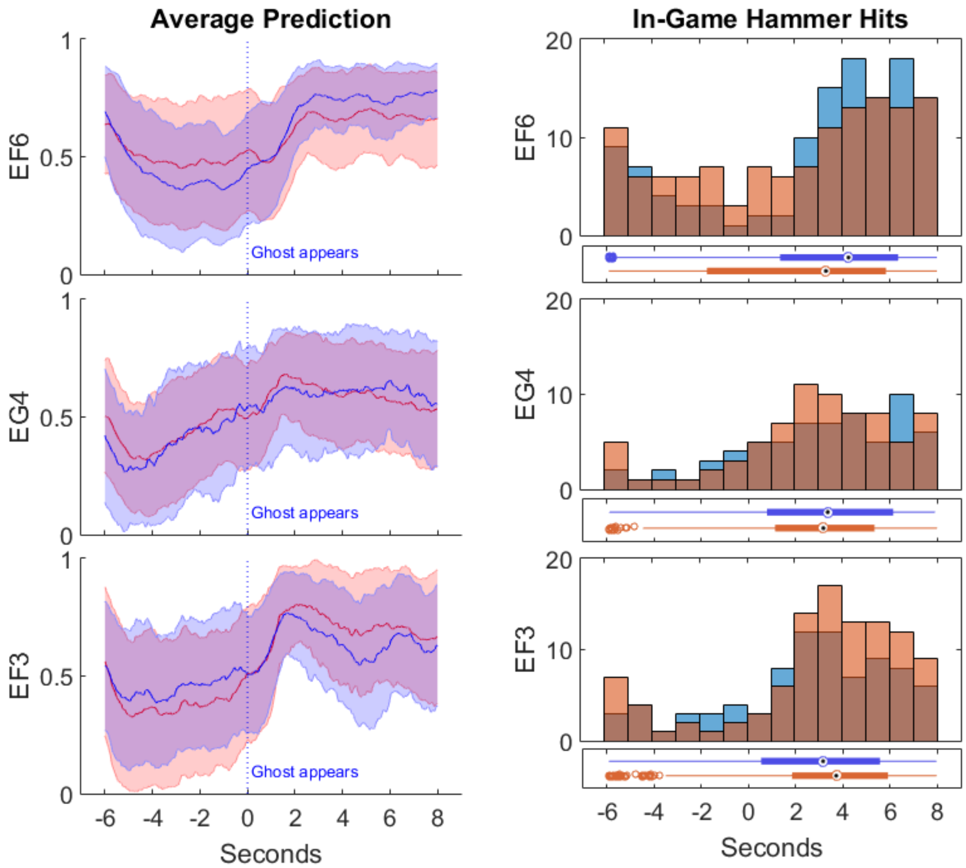
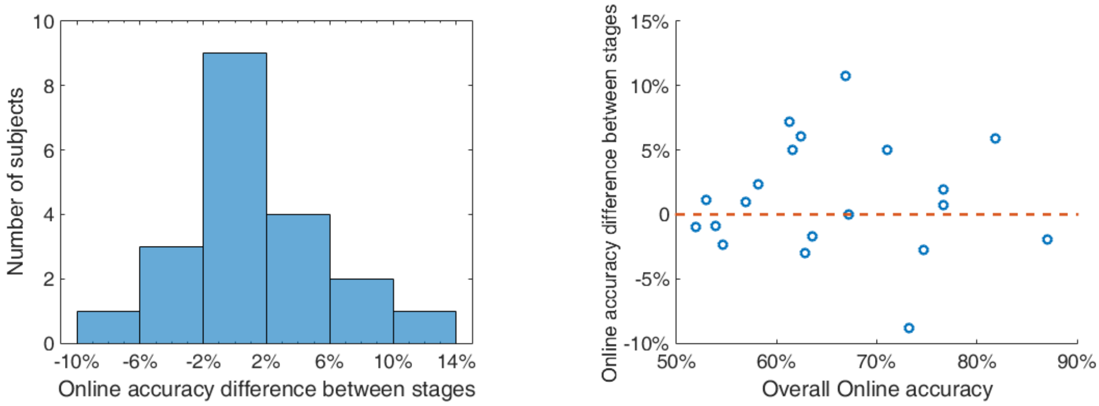


FIGURE 3.5: Comparison of different stages performance for subjects EF6(top), EG4(center) and EF3 (bottom). Average online prediction (left). Histogram and boxplot of hammer strikes (right). The curves on the left show the average ± standard deviation (highlighted area) sLDA prediction *p_MI* curves. The graphs on the right show hammer strikes per round. The timing on the x axes is relative to the appearance of the ghost. The orange curves, boxplots and histograms refer to the adaptive stage, the blue equivalents refer to the non-adaptive stage.



(A) (B)

FIGURE 3.6: Histogram (A) of the differences in performance from adaptive to non-adaptive stage and scatter plot (B) of the same difference as a function of user online accuracy.

3.3 Evolution between sessions

Four out of the twenty subjects did the experiment multiple times. EF7, EF8 and EG4 did 2 sessions and subject EE3 did 3 sessions. Table 3.3 shows the difference in performance between sessions for these 4 subjects.

TABLE 3.3: Difference in performance between sessions for subjects that volunteered multiple times.

Subject code	Online Accuracy (%)	△ Online Accuracy (%)	△Trial Accuracy(%)	△ Stars per Round	△ Control
EE3	63.6	1.7	4.6	5	1
EE3*	65.3	7.0	3.8	41	1.5
EF7	62.4	13.6	19.8	57	1.5
EF8	61.6	-4.0	-7.4	-14	1.5
EG4	67.2	5.1	5.6	-14	-0.5
Mean	64.0	4.7	5.3	15	1.0
Std	2.2	6.5	9.6	32	0.9

* from second to the third session

The differences in Table 3.3 are visualized in Figure 3.7 that shows the difference in performance between stage and session for each user.

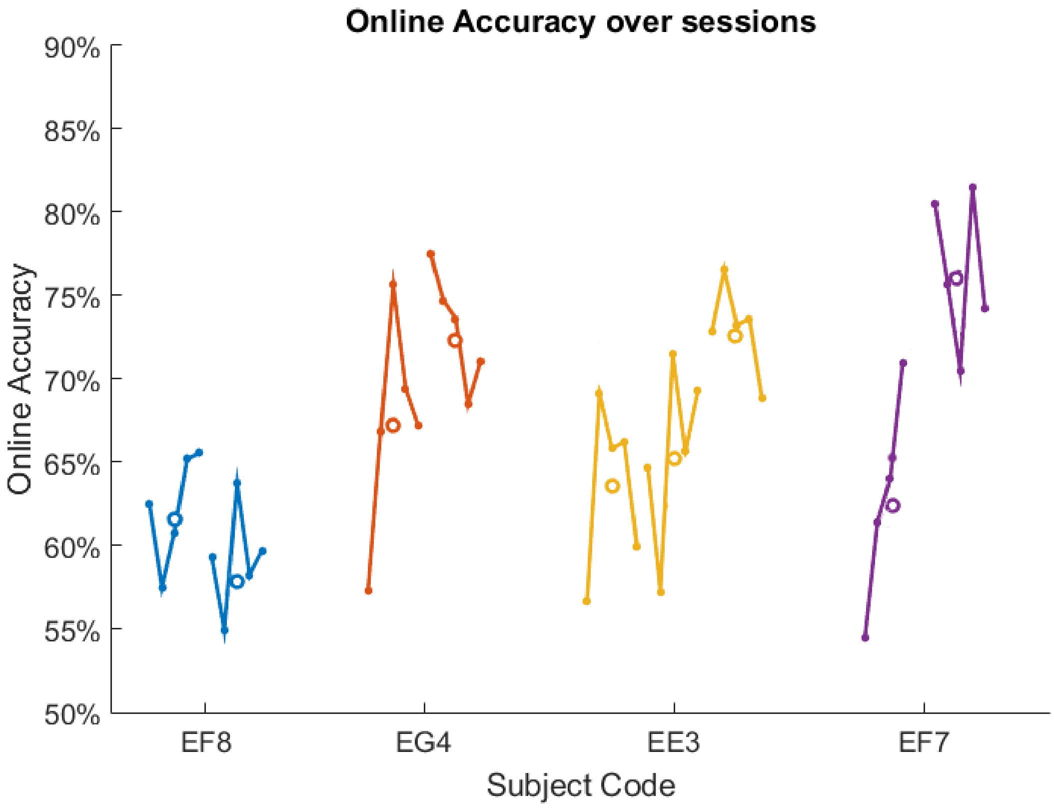


FIGURE 3.7: Subjects' Online Accuracy between sessions. Each color represents a subject. The lines and dots show the performance in each round, the white centered circle is the session's average online accuracy. For each subject, first sessions are on the left, last sessions are on the right.

3.4 Random Performance

Random performance level was assessed by permutation tests. Each recording had its EEG features shuffled in intervals from 1/8 to 5 seconds. This process was repeated several times to create 10000 random sessions (2×10^6 trials). The average accuracy of the "random performers" is 50.7±1%. The best 5% of the simulations had accuracy over 52.6% and the top percentile had accuracy over 53.4%. The accuracy of the top percentile was used as the threshold of minimum control, so users with accuracy under this value were considered not in control. The results of all simulations are shown in Figure 3.8.

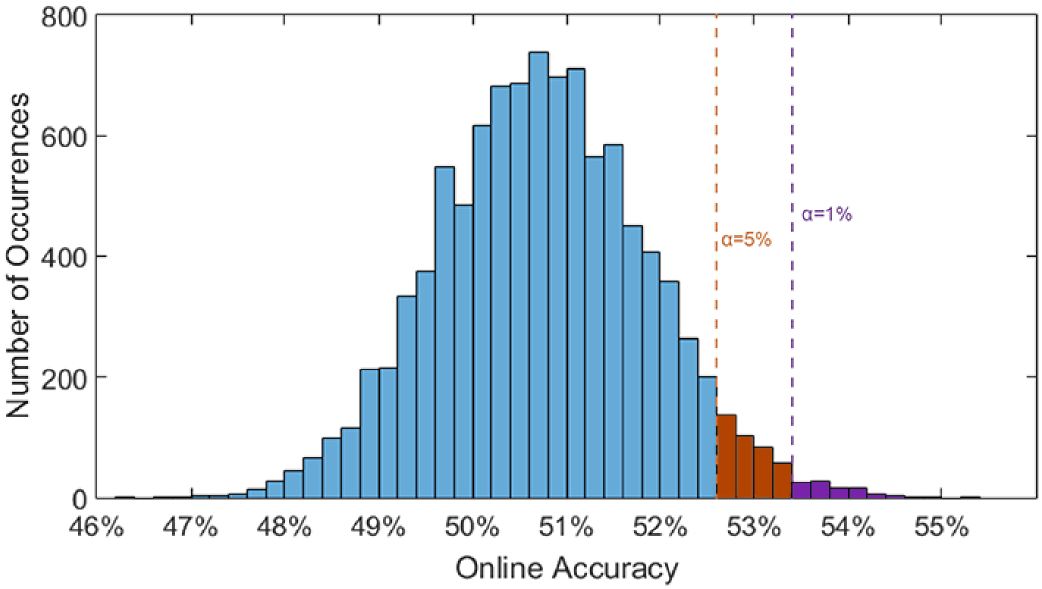


FIGURE 3.8: Histogram with the Online Accuracy from 10000 random performers.

4 Discussion

4.1 Results analysis

The aim of this study was to assess the effectivity of online co-adaptive training paradigms. This results with 20 naive users suggest that fully-automatic online coadaptation allows users to gain meaningful levels of non-adaptive BCI control after 30 minutes of co-adaptive training.

4.1.1 Overall system

The bottom panel in Figure 3.1 shows results of user EF2 who achieved Online Accuracy of 51.9%. Prediction probability, histogram of hammer strikes and boxplot of the timing of the strikes suggest that the user had no control. The upper panel shows results for user ED6, who achieved the highest Online Accuracy = 87%. The curves show a clear increase in the predicted probability when the ghost was visible on the screen. Hammer strikes are clustered around the same period with a short delay. The middle panel shows results of user EF1 with Online Accuracy = 62.9%. The curves show an increase in predicted probability after the ghost appears. Hammer strikes increase 3 seconds after the ghost's appearance (peak at second 4) until 1 second after it disappears. Such a curve suggests that there is a tendency to have control. According to the permutation tests, users with Online Accuracy over 53.4% gained levels of BCI control significantly better than random. Being so, 18 out of 20 users can be considered as being in control.

Studies with similar BCI systems [19], [21]–[24] were tested with the system constantly adapting. This studies had a subject number ranging from 4 up to 22, with 75% up to 100% of the subject being able to perform better than random. These further confirms that our results are within range of what was expected.

Analyzing Table 3.1 results seem more favorable if we consider the accuracy over 6 seconds instead of 1. The average online accuracy is 66,4% and the average trial accuracy is 74.7% . This suggests that online fast responsive systems sacrifice accuracy for response speed. Meaning that systems with higher accuracy can be created for tasks that do not require fast feedback.

The strong correlation between true positive hammer strikes, game score and BCI online accuracy indicate that the feedback system was working as intended.

Figure 3.4.A shows the relation between self-reported control and BCI online accuracy. This relation provides more evidence that subjects with better performances were getting positive feedback and felt more in control. Figure 3.4.B shows there is also a relation between satisfaction and online BCI accuracy, but in this case, online accuracy explains less of the satisfaction variability (lower adjusted R Squared) and satisfaction has a higher interception value. This mean that its possible to make an enjoyable game even for users with a non-ideal level of control. This is of great necessity to engage subjects in this kind of experiment, and it is especially important for experiments where subjects need to be motivated to volunteer multiple times. In this experiment 19 out of 20 users reported to be interested in volunteering for a second session.

4.1.2 Adaptive vs non-adaptive system

The study of the performance in the non-adaptive stage of the game was of key importance for this study. If a user intends to use a BCI system in a real-life scenario it is not expected that the BCI "knows" the right answer from the start, so it cannot adapt. This study aimed to prove that a BCI that was calibrated using online interactive co-adaption can maintain his performance once the calibration phase is over.

Using both parametric (t-test) and non-parametric tests (wilcoxon test) there is no evidence to suggest that there is a significant difference in performance between the co-adaptive stage and the non-adaptive stage. T-test suggests that the true shift in performance between stages is between -2% and +3.4% (α=0.01). There were a few exceptions to this rule, the subjects with codes EF3 and EF6 had a significant (higher than 2 times the standard deviation value) decrease and increase in performance between stages. Subject EF3 reported to be increasingly tired both in the user experience form and informally in the pauses between rounds, this may explain a gradual decrease in accuracy over time. Subject EF6 made a remarkable comment in the pause between the second and the third round of the experiment, he reported: "I think I now understood how to do it". This comment preceded a high increase in accuracy that lasted until the end of the second stage. In the end of the experiment subject EF6 reported to change mental strategy between the first and second round from imagining hammering the ghost to imagining squeezing it. The user reported to start performing better in the end of the second round. This highlights the importance of properly instructing user on how to use BCI as well as the importance of making learning environments where users feel engaged and eager to learn and try new strategies - in this case, to win the game!

4.1.3 Subject performance between sessions

The main goal of this project was to compare the performance of the BCI in the adaptive and non-adaptive stage for naïve users. However, some subjects volunteered to repeat the experiment multiple times. This was an opportunity to try to replicate, in a smaller scale, some results observed in 2012 by Faller et al. [21] where 9 out of 12 co-adaptive BCI users improve their performance between sessions. In this experiment 3 out of the 4 subjects improved their accuracy between sessions, this includes a subject that improved both from the first to second and from second to third session. These results are preliminary, and the number of subjects is not enough to assert a strong conclusion, but the data suggests that there may be an increase in online accuracy as users get more experienced with BCI. This would come as no surprise as it has been shown several times that BCI users benefit from experience [34]–[36].

4.2 Future Work

Due to time constraints, many promising ideas to improve this project were not implemented, some never left the drawing board, and some were fully implemented but not tested.

4.2.1 Error Related Potentials detection

Error-related Potentials (ErrPs), are the alterations of the EEG traces related to the subject perception of erroneous responses. The detection of ErrPs may be a strong addition to an ERD-based BCI system since it detects what the user perceives as misclassifications. In previous research, ErrPs were detected online using time [37] and frequency [38] analysis.

In our experiment, the user controlled a hammer that has 3 possible states: rest, charge and strike. The charging state is the intermediate state between resting and striking. When the ghost is not present on the screen, charging precedes an erroneous strike. The users should be able to perceive this error and avoid it. An improved version of our BCI system would detect an ErrP that would emerge when the hammer goes to charging state in time to avoid the strike.

Figure 4.1 shows the difference in time domain after the hammer charges while the ghost in on the screen (intentional charge) and while it is not (unintentional charge).

4.2.2 The FORCe

Artifact rejection is made only during sLDA training because online rejection of a noisy signal would require giving incorrect feedback to the user every time there is an artifact. To fix this problem we attempted to do artifact reduction with the *Fully Online and Automated Artifact Removal for Brain-Computer Interfacing* (FORCe) [39]. This method uses independent component analysis (ICA) to identify artifacts and output the clean signal. To use this tool at its full potential, we developed the whole BCI system in python, implemented the FORCe and adapted the game for Linux. Unfortunately, the fastest feedback we could get from the system would come with a consistent delay from 1.5 to 2.5 seconds. This idea was abandoned has responsiveness of the online feedback was prioritized.

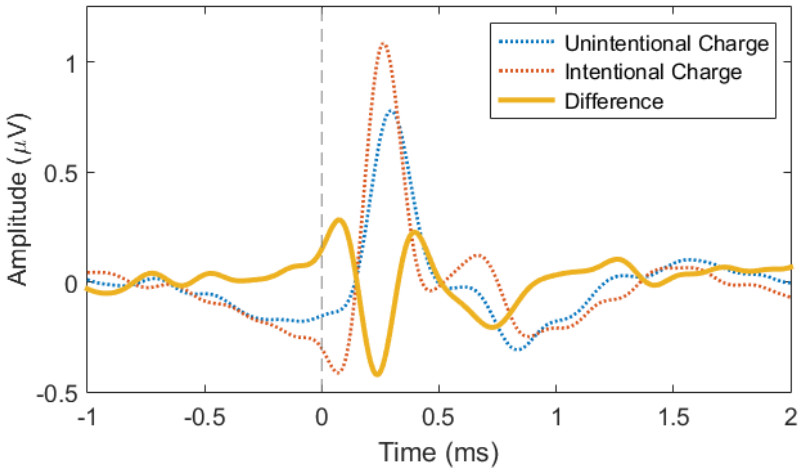


FIGURE 4.1: Grand average EEG (Cz) from the 25 recordings, showing the signal from 1 second before to 2 seconds after the start of the hammer charging animation. Intentional and unintentional charges are compared, and the difference is plotted in yellow.

4.2.3 Multiplayer improved gameplay

It was mentioned before that an engaging and challenging gameplay was of key importance for the users to feel eager to learn. Although this game is an improvement relative to many standard MI paradigms, it is still quite repetitive. In 2013, Bonnet et al. [40] created two multiplayer BCI games where users could either cooperate or compete with each other. This study, reports similar user performance for solo, collaborative and competitive mode, but users reported to have more fun and to feel more motivated in the multiplayer modes. This inspired us to make a competitive version of the game. This feature was fully implemented, but not in time for consistent testing and result analysis.

Furthermore, it would be interesting to create different challenges for each round (whack-a-mole, tennis match, obstacle run) and changing the scoring system to promote competitiveness and eagerness to improve each round.

5 Conclusion

The presented system was designed to train the user and the BCI to be ready for use in 30 minutes. 18 out of 20 naïve users successfully achieved control with the co-adapted BCI system and were able to play the game.

The main requirement for this system was that its accuracy would not decrease after the co-adaptation stops. Only 1 out of the 18 subjects that were considered to be in control had a significant decrease in online accuracy from the adaptive stage to the non-adaptive. This decrease (from 74.6% to 65.8%) was not enough to make the user drop below the chance level. This indicates that after 30 minutes of user training and BCI calibration the system is ready to use.

These promising results indicate that this kind of BCI system does not rely on the constant adaptation to work. This means that systems like this may be implemented in real life scenarios, where the system does not have constant access to the user's intention. This kind of system with few degrees of freedom can be used to control switches and wheelchairs which may be life changing for people with severe motor disabilities.

Bibliography

[1] P. Tate, "Central and peripheral nervous system", in *Seeley's principles of anatomy and physiology.* McGraw-Hill, 2012, pp. 299 –342.

[2] ——, "Integration of nervous system functions", in *Seeley's principles of anatomy and physiology.* McGraw-Hill, 2012, pp. 343 –368.

[3] G. J. Mogenson, D. L. Jones, and C. Y. Yim, "From motivation to action: functional interface between the limbic system and the motor system.", *Progress in Neurobiology*, vol. 14, no. 2-3, pp. 69–97, 1980, ISSN: 0301-0082.

[4] Y. Kajikawa and C. E. Schroeder, "How local is the local field potential?", *Neuron*, vol. 72, no. 5, pp. 847–858, 2011, ISSN: 0896-6273.

[5] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review.", *Sensors (Basel, Switzerland)*, vol. 12, no. 2, pp. 1211–79, 2012, ISSN: 1424-8220.

[6] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control.", *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–91, 2002, ISSN: 1388-2457.

[7] J. D. R. Millán, R Rupp, G. R. Müller-Putz, R Murray-Smith, C Giugliemma, M Tangermann, C Vidaurre, F Cincotti, A Kübler, R Leeb, C Neuper, K.-R. Müller, and D Mattia, "Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges.", *Frontiers in Neuroscience*, vol. 4, 2010, ISSN: 1662-453X.

[8] G. Pfurtscheller, B. Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, T. O. Zander, G. Mueller-Putz, C. Neuper, and N. Birbaumer, "The hybrid bci.", *Frontiers in Neuroscience*, vol. 4, p. 30, 2010, ISSN: 1662-453X.

[9] T. O. Zander and C. Kothe, "Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general", *Journal of Neural Engineering*, vol. 8, no. 2, p. 025 005, 2011, ISSN: 1741-2560.

[10] D. Marshall, D. Coyle, S. Wilson, and M. Callaghan, "Games, gameplay, and bci: the state of the art", *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 5, no. 2, pp. 82–99, 2013, ISSN: 1943-068X.

[11] L. A. Farwell and E Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials.", *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–23, 1988, ISSN: 0013-4694.

[12] M. A. Wenzel, I. Almeida, and B. Blankertz, "Is neural activity detected by erp-based brain-computer interfaces task specific?", *PLOS ONE*, vol. 11, no. 10, D. Marinazzo, Ed., e0165556, 2016, ISSN: 1932-6203.

[13] S. T. Morgan, J. C. Hansen, and S. A. Hillyard, "Selective attention to stimulus location modulates the steady-state visual evoked potential.", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 10, pp. 4770–4, 1996, ISSN: 0027-8424.

[14] C. Neuper, R. Scherer, M. Reiner, and G. Pfurtscheller, "Imagery of motor actions: differential effects of kinesthetic and visual–motor mode of imagery in single-trial eeg", *Cognitive Brain Research*, vol. 25, no. 3, pp. 668–677, 2005, ISSN: 09266410.

[15] G Pfurtscheller, P Linortner, R Winkler, G Korisek, and G Müller-Putz, "Discrimination of motor imagery-induced eeg patterns in patients with complete spinal cord injury.", *Computational Intelligence and Neuroscience*, vol. 2009, p. 104 180, 2009, ISSN: 1687-5273.

[16] R. Scherer, A. Schwarz, G. R. Muller-Putz, V. Pammer-Schindler, and M. L. Garcia, "Game-based bci training: interactive design for individuals with cerebral palsy", in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, 2015, pp. 3175–3180, ISBN: 978-1-4799-8697-2.

[17] F. Lotte, F. Larrue, and C. Mühl, "Flaws in current human training protocols for spontaneous brain-computer interfaces: lessons learned from instructional design.", *Frontiers in Human Neuroscience*, vol. 7, p. 568, 2013, ISSN: 1662-5161.

[18] D. J. McFarland, W. A. Sarnacki, T. M. Vaughan, and J. R. Wolpaw, "Brain-computer interface (bci) operation: signal and noise during early training sessions", *Clinical Neurophysiology*, vol. 116, no. 1, pp. 56–62, 2005, ISSN: 13882457.

[19] J. Faller, R. Scherer, U. Costa, E. Opisso, J. Medina, and G. R. Müller-Putz, "A co-adaptive brain-computer interface for end users with severe motor impairment.", *PloS one*, vol. 9, no. 7, e101168, 2014, ISSN: 1932-6203.

[20] M Gomez-Rodriguez, J Peters, J Hill, B Schölkopf, A Gharabaghi, and M Grosse-Wentrup, "Closing the sensorimotor loop: haptic feedback facilitates decoding of motor imagery", *Journal of Neural Engineering*, vol. 8, no. 3, p. 036 005, 2011, ISSN: 1741-2560.

[21] J. Faller, C. Vidaurre, T. Solis-Escalante, C. Neuper, and R. Scherer, "Autocalibration and recurrent adaptation: towards a plug and play online erd-bci", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 3, pp. 313–319, 2012, ISSN: 1534-4320.

[22] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, "Co-adaptive calibration to improve bci efficiency", *Journal of Neural Engineering*, vol. 8, no. 2, p. 025 009, 2011, ISSN: 1741-2560.

[23] A. Schwarz, R. Scherer, D. Steyrl, J. Faller, and G. R. Muller-Putz, "A co-adaptive sensory motor rhythms brain-computer interface based on common spatial patterns and random forest", in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2015, pp. 1049–1052, ISBN: 978-1-4244-9271-8.

[24] R. J. Kobler and R. Scherer, “Restricted boltzmann machines in sensory motor rhythm brain-computer interfacing: a study on inter-subject transfer and co-adaptation”, in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, 2016, pp. 000 469–000 474, ISBN: 978-1-5090-1897-0.

[25] C. Jeunet, E. Jahanpour, and F. Lotte, “Why standard brain-computer interface (bci) training protocols should be changed: an experimental study”, *Journal of Neural Engineering*, vol. 13, no. 3, p. 036 024, 2016, ISSN: 1741-2560.

[26] S. Kober, M. Ninaus, E. V. C. Friedrich, and R. Scherer, “Bci and games: playful, experience-oriented learning by vivid feedback?”, in *Brain-computer interfaces handbook : Technological and theoretical advances*, C. S. Nam, A. Nijholt, and F. Lotte, Eds., CRC Press Taylor & Francis Group, 2018, p. 207.

[27] J. Cunha and R. Scherer, “Are online co-adaptive sensorimotor rhythm brain-computer interface training paradigms effective?”, in *2018 International Conference on Cyberworlds*, 2018.

[28] C. Lopes Dias, A. I. Sburlea, and G. R. Müller-Putz, “Masked and unmasked error-related potentials during continuous control and feedback”, *Journal of Neural Engineering*, vol. 15, no. 3, p. 036 031, 2018, ISSN: 1741-2560.

[29] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, “Single-trial analysis and classification of erp components — a tutorial”, *NeuroImage*, vol. 56, no. 2, pp. 814–825, 2011, ISSN: 1053-8119.

[30] R. A. Bjørklund, “Reaction time and movement time measured in a key-press and a key-release condition”, *Perceptual and Motor Skills*, vol. 72, no. 2, pp. 663–673, 1991, ISSN: 0031-5125.

[31] W. B. W. Daud and R. Sudirman, “Time frequency analysis of electrooculograph (eog) signal of eye movement potentials based on wavelet energy distribution”, in *2011 Fifth Asia Modelling Symposium*, IEEE, 2011, pp. 81–86, ISBN: 978-1-4577-0193-1.

[32] S. D. Muthukumaraswamy, “High-frequency brain activity and muscle artifacts in meg/eeg: a review and recommendations”, *Frontiers in Human Neuroscience*, vol. 7, p. 138, 2013, ISSN: 1662-5161.

[33] J. Cohen, “A coefficient of agreement for nominal scales”, *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, 1960, ISSN: 0013-1644.

[34] R. Leeb, D. Friedman, G. R. Müller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller, “Self-paced (asynchronous) bci control of a wheelchair in virtual environments: a case study with a tetraplegic.”, *Computational Intelligence and Neuroscience*, vol. 2007, p. 79 642, 2007, ISSN: 1687-5265.

[35] R. Leeb, S. Perdikis, L. Tonin, A. Biasiucci, M. Tavella, M. Creatura, A. Molina, A. Al-Khodairy, T. Carlson, and J. d.R. Millán, “Transferring brain–computer interfaces beyond the laboratory: successful application control for motor-disabled users”, *Artificial Intelligence in Medicine*, vol. 59, no. 2, pp. 121–132, 2013, ISSN: 0933-3657.

[36] M. Rohm, M. Schneiders, C. Müller, A. Kreilinger, V. Kaiser, and G. R. Müller-Putz, “Hybrid brain–computer interfaces and hybrid neuroprostheses for restoration of upper limb functions in

individuals with high-level spinal cord injury", *Artificial Intelligence in Medicine*, vol. 59, no. 2, pp. 133–142, 2013, ISSN: 0933-3657.

[37] B. Dal Seno, M. Matteucci, and L. Mainardi, "Online detection of p300 and error potentials in a bci speller", *Computational Intelligence and Neuroscience*, vol. 2010, pp. 1–5, 2010, ISSN: 1687-5265.

[38] S. Rousseau, C. Jutten, and M. Congedo, *The error-related potential and BCIs*. 2012.

[39] I. Daly, R. Scherer, M. Billinger, and G. Muller-Putz, "Force: fully online and automated artifact removal for brain-computer interfacing", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 5, pp. 725–736, 2015, ISSN: 1534-4320.

[40] L. Bonnet, F. Lotte, and A. Lécuyer, "Two brains, one game: design and evaluation of a multi-user bci video game based on motor imagery", *IEEE Transactions on Computational Intelligence and AI in games*, vol. 5, no. 2, pp. 185–198, 2013.

A Appendix - Study Information Sheet

Study information sheet

1. Explanation of the recorded signals
2. Goal of the experiment
3. General schedule
4. The game
5. The game rounds - What is going on?
6. Which movements should be imagined
7. Code of behavior during the experiment
8. Time for questions

1. Dear participant

Within the next 1.5 hours, you will undergo a BCI (Brain- Computer Interface) game. During the whole experiment, we will record your brain waves (EEG – Electroencephalography).

2. Goal of the game

Mental activity, the imagination of specific movements of different body parts, leads to measurable changes in your brain activity.
The goal of this training is to discriminate the imagination of hand movement from relaxation on the basis of your EEG. The final goal is to use this technology to help people with severe disability to regain autonomy.

3. General schedule

The experiment is divided into the following blocks.

Explanation	Electrode Mounting	Experiment								
		Round 1	Break	Round 2	Break	Round 3	Break	Round 4	Break	Round 5

We have already started with the explanation block during which you get to know about the experiment. As we mount the electrodes on your head, feel free to ask any questions that may occur. After that, the actual experiment starts.

4. The game

One game is split into 5 short round. Each lasts for about 10 minutes. Within each round, a cartoonish ghost will appear and disappear intermittently. Your task is to imagine hand movement while the ghost is on the screen. The hand you choose to imagine should be the one that feels more natural and vivid to imagine. You should **never change hands** after you picked the one that you will imagine!

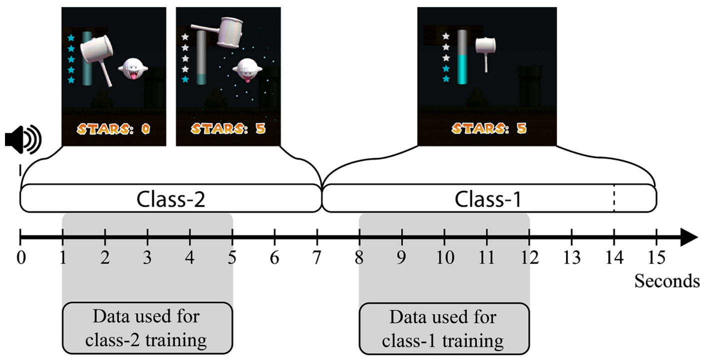
In the game you control an hammer that will slam the ghost. Once you start performing imaginary, the hammer should start charging, and if you manage to hold it for a few seconds you will slam the ghost, and win points equivalent to your energy stars!

You recover energy by resting while the ghost is not on the screen.

Charging Hitting Resting

5. The 5 Rounds: What is going on?

Since every person had a unique set of brain signals our system tries to learn from scratch what happens in your brain during your imagination. During the first 3 rounds, our system will be adapting to the patterns that it recognizes in your brain. Therefore it's extremely important that you **keep your imagination technique as constant as possible within each round** so that there is a stable pattern. In the last 2 rounds, the game will be running based on the 3rd round.



6. Which movement should be imagined?

Here are some examples:

- Making a fist,
- Pressing an anti-stress ball,
- Scratching your thumb with your index finger,
- Squeezing fruit

Before you start, Execute different movements for both your hands, then imagine the movements you executed before and select the movement and hand which you can imagine best in terms of vividness, concentration, repeatability.

Remember:

- Start the imagination of the movement as soon as the ghost appears.
- Relax and avoid thinking about your hands as soon as the ghost is gone.
- Focus on kinesthetic/tactile sensation (put attention on what you (would) feel your hand).
- You found the right one if you can easily recall the mental task and it is life-like in your brain.

7. Code of behavior

•Try to sit comfortable and avoid movements during the measurement.
•Put your hands comfortably on your lap/or on the table.
•Relax the muscles of your face, neck, shoulders and lower jaw!
•Please avoid blinks and swallowing in a repetitive pattern (e.g. blinking every time the ghost appears)

8. If there are any questions left please ask the conductor now!

B Appendix - User Experience Form

First Stage (Round 1 to 3)

During the game how energetic where you?
Tired 1 2 3 4 5 6 7 8 9 10 Alert
Did you feel in control of the game?
Very little 1 2 3 4 5 6 7 8 9 10 Very much
How do you classify your performance over time in the game ?
Decreasing 1 2 3 4 5 6 7 8 9 10 Increasing
Overall satisfaction with the system?
Unsatisfied 1 2 3 4 5 6 7 8 9 10 Satisfied

Second Stage (Round 4 and 5)

During the second stage how energetic where you?
Tired 1 2 3 4 5 6 7 8 9 10 Alert
During the second stage did you feel in control of the game?
Very little 1 2 3 4 5 6 7 8 9 10 Very much
How do you classify your performance over time in the game during the second stage ?
Decreasing 1 2 3 4 5 6 7 8 9 10 Increasing
Overall satisfaction with the system in the second stage:
Unsatisfied 1 2 3 4 5 6 7 8 9 10 Satisfied

Overall

How much physically effort or activity was required?
Low 1 2 3 4 5 6 7 8 9 10 High
How much mental effort or concentration was required?
Low 1 2 3 4 5 6 7 8 9 10 High
What was the level of frustration during the task?
Low 1 2 3 4 5 6 7 8 9 10 High

After using the system, do you now feel headache pain?
Not at all 1 2 3 4 5 6 7 8 9 10 Extremely
After using the system, your eyes feel tired, strained or stressed?
Not at all 1 2 3 4 5 6 7 8 9 10 Extremely
After using the system, you muscles feel tired, strained or stressed?
Not at all 1 2 3 4 5 6 7 8 9 10 Extremely
Overall this was physically :
Comfortable 1 2 3 4 5 6 7 8 9 10 Uncomfortable

Is this the first time you use a brain-controlled interface? Yes☐ No☐
Would you be interested in repeating this experiment during the next month? Yes☐ No☐