NEUROMODULATION FOR SUPPORTING LOWER LIMB MOVEMENT REHABILITATION

Inês Silva Veríssimo

MESTRADO INTEGRADO EM ENGENHARIA BIOMÉDICA E BIOFÍSICA
Especialização em Sinais e Imagens Médicas

Dissertação orientada por:
Prof. Doutor Alexandre da Rocha Freire de Andrade
e Prof. Doutor Jason Farquhar

2018
"Again, of all the things that come to us by nature we first acquire the potentiality and later exhibit the activity (...) For the things we have to learn before we can do them, we learn by doing them”

— Aristotle, *Nicomachean Ethics*
Acknowledgements

When thinking about this section of my thesis, I could not avoid remembering a speech given by Dr. António Sampaio da Nóvoa, where he mentions the three levels of gratitude as defined by St. Thomas of Aquina. I will not reproduce Dr. Nóvoa’s speech, but simply give a brief definition of the levels that one can reach when giving acknowledgments. The superficial level is the intellectual recognition of the benefit given (thank you), the intermediate level is when we give grace for what we received (gracias) and, finally, the most profound level of gratitude is when we are bond to someone, i.e., when we are committed to return the kindness that was given to us. The latter is inherently expressed by the portuguese word obrigado. Thus, I give my most sincere and profound gratitude to every person/entity that, in one way or another, contributed for the development of this dissertation.

First of all, to the Erasmus+ program for enabling this experience, which allowed me to grow both personally an professionally.

Obrigada to my internal supervisor, Professor Alexandre Andrade, for the constant support and feedback given throughout this academic journey.

To Dr. Jason Farquhar, obrigada for the guidance and help in the development of this project. For teaching me so much about BCIs and neuroscience, and showing me that we should always question everything and try to do the best that we can.

To the members of the Donders BCI group, for receiving me and making me feel as part of the team. A special obrigada to Ceci Verbaarschot that, without having to, also acted as my supervisor. Obrigada for all the patience when answering my infinite questions and helping me with the practical issues of my project. Obrigada to Anne Gerrits and Yvonne Visser, for the friendship and for being "guinea-pigs" in all my pilot experiments.

Obrigada to my family and friends, both from my home-country and abroad, without whom this experience would have been considerably more difficult. A special mention to my friends Isabel Barradas, Iara Almeida, and Joana Boita (a much needed familiar face in a strange country). Obrigada for all the support, the countless hours of pep talks, and for making my time as a university student so much more fun.

At last, but definitely not least, to my mother. Obrigada por todo o amor e apoio que me deste ao longo da minha vida. Por me ensinares a lutar pelos meus objectivos sem medos e me mostrares que, com empenho e persistência, tudo se consegue.
Abstract

Stroke is a devastating incident, affecting more than fifteen million people worldwide each year. Most stroke patients have motor impairments that often result in disability, such as impaired gait and inability to perform day-to-day activities. When questioned, stroke survivors report that their primary wish is to regain their lost independence. For this reason, much of the standard rehabilitation strategies employed involve the repetition of the lost motor skill (e.g.: re-training of walking), to promote motor learning. If a patient is not able to perform the movement himself, the movement is assisted by a physical therapist that will move their limbs for them. This type of rehabilitation is a bottom-up approach – limb movement influencing brain activation patterns and, thus, encouraging brain plasticity. In contrast, our main goal is to study the effects of BCI training, a top-down approach, as a tool to promote neuromodulation and, ultimately, motor learning.

The goal of this dissertation is to study the effectiveness of an electroencephalography (EEG)-based Brain-Computer Interface (BCI) robot training protocol on healthy participants, in the acquisition of a new motor skill: toe abduction. For this, twenty-one subjects were divided into three different conditions - active robotic feedback, visual feedback and mock robotic feedback - and performance was assessed throughout five experimental training sessions.

Results show that participants that were subjected to contingent BCI feedback, through the robotic apparatus, had a significant increase between sessions in muscle activation and in the range of motion (ROM), our primary outcome to access the ability to abduct the big toe. However, there were no significant differences in these outcome measurements, when performing an inter-condition comparison. Additionally, no significant results were found when comparing changes in the cortical oscillatory activity, as a measure of brain plasticity. These results indicate that, although participants from the active robotic condition showed the desired behavior, the overall evidence obtained is not sufficient to state that this type of protocol is more efficient in the promotion of motor learning of a novel skill in healthy participants.

We propose an innovative motor learning approach consisting of enhancing brain activation using a brain-controlled robot for lower limb training, but future work needs to be developed in order to confirm the effectiveness of this approach when compared to conventional techniques.

Keywords: Brain-Computer Interface, Electroencephalography, Motor Learning, Robotic Feedback, Brain Plasticity, Neuromodulation, Rehabilitation, Stroke
Resumo

Vários estudos sugerem que o uso de neurofeedback pode melhorar o desempenho físico e/ou cognitivo humano. Este conceito pode ser definido num espectro contínuo: num extremo temos o comportamento disfuncional, e no extremo oposto o comportamento ótimo. Desta forma, as alterações ao desempenho introduzidas pelo neurofeedback podem advir de situações onde indivíduos que sofrem de algum gênero de incapacidade física ou cognitiva recuperam, atingindo um comportamento normativo, ou de situações onde os indivíduos que, a priori, demonstram comportamentos padrão, desenvolvem as suas capacidades, atingindo um comportamento perto do ótimo. Ainda que este último caso seja um objetivo tangível da investigação contemporânea, com cada vez mais estudos a focarem-se no uso de ferramentas de neuropotenciação para promover determinadas capacidades cognitivas (como a velocidade de processamento, memória de trabalho ou funções executivas) em profissionais que estão sujeitos a grandes cargas de trabalho ou a estudantes com dificuldades de aprendizagem ou concentração, o primeiro tópico foi o que deu acento ao desenvolvimento desta dissertação.

Acidentes vasculares cerebrais (AVC) são uma das principais causas de mortalidade em países com elevadas taxas de desenvolvimento económico e social, afetando anualmente mais de quinze milhões de pessoas a nível mundial. São caracterizados por problemas na irrigação sanguínea do cérebro, causando a morte celular de neurônios especializados, o que leva a que determinadas regiões do cérebro deixem de funcionar devidamente. O AVC geralmente causa um impacto significativo na vida funcional, cognitiva e social do paciente, dado que as repercussões muitas vezes levam a incapacidades crónicas. O surgimento de deficiências motoras dos membros superiores e inferiores são extremamente comuns, o que leva a que grande parte das estratégias de reabilitação para pacientes que sofreram um AVC se concentrem na recuperação da mobilidade.

As técnicas de reabilitação convencionais podem ser definidas como abordagens "de baixo para cima": o tratamento foca-se na repetição intensiva de movimentos no membro afetado, com o objetivo de influenciar o sistema neuronal e a sua plasticidade. Contudo, este gênero de terapêuticas apenas beneficia certa de 50% dos pacientes, o que significa que a outra metade da população que recorre a esta abordagem continua parcialmente dependente de terceiros para a realização das suas atividades diárias. Protocolos que recorrem aos princípios da aprendizagem motora têm tentado melhorar este quadro. A aprendizagem motora está associada à proliferação de dendrites e à formação de novas ou alteração de sinapses pré-existentes. Desta forma, este tipo de aprendizagem pode ser uma ferramenta útil para induzir a reorganização cortical necessária à reaprendizagem de tarefas motoras.

As Interfaces Cérebro-Máquina (BCI, do inglês Brain-Computer Interface) são sistemas que fornecem uma alternativa às vias de comunicação comuns, pois criam uma linha direta de comunicação e controlo entre o cérebro e aparelhos físicos ao traduzir, em tempo-real, os diferentes padrões de atividade cerebral em comandos. Protocolos de reabilitação que recorram ao uso de BCIs para proporcionar feedback multissensorial podem, potencialmente, modular a atividade cerebral do utilizador, induzindo alterações plásticas ao hemisfério lesionado e, assim, promover o restauro das funções motoras afeta-
das depois do AVC. Este gênero de sistema pode ser, então, visto como uma abordagem "de cima para baixo" à neuroreabilitação: neste caso, estimulamos a plasticidade cerebral ao incentivar a participação ativa do paciente no processo de reabilitação, de forma a produzir o movimento desejado, fechando, assim, o ciclo sensoriomotor.

A presente tese teve como objetivo o estudo de um protocolo de treino BCI, com base na aquisição de sinais eletroencefalográficos em tempo-real, acoplado a um aparelho robótico. Vinte e um sujeitos saudáveis participaram nas cinco sessões experimentais deste estudo, sendo divididos em três grupos: feedback robótico, feedback visual e o grupo de controlo, que recebeu feedback robótico falso. Pretendíamos utilizar este sistema para ensinar os participantes a efetuar um movimento que não dominavam a priori, a abdução do hálux (dedo grande do pé) direito. Para pacientes que sofreram um AVC, o processo de recuperação é, muitas vezes, demorado e frustrante. Por isso, quisemos criar uma tarefa que recriasse, até certo ponto, este tipo de obstáculos à aprendizagem motora em sujeitos saudáveis. O sistema muscular humano apresenta músculos funcionais que permitem a realização deste movimento (abductor hallucis), porém a maioria da nossa amostra de sujeitos não conseguia realizar o movimento antes de participar na experiência.

Os resultados mostraram que os participantes que receberam feedback robótico, devido ao controlo ativo do sistema BCI, tiveram um aumento significativo da atividade muscular assim como do alcance do movimento (medido em graus), ao longo das sessões. Contudo, ao comparar estes resultados entre os vários grupos, não foram observadas diferenças significativas. Alterações na atividade oscilatória do córtex motor também foram analisadas, através do cálculo do ERD (do inglês, event-related desynchronization, marcador que indica a supressão das ondas beta no córtex motor) e das percentagens de classificação obtidas. Mas, mais uma vez, não foram observadas diferenças significativas entre grupos ou sessões. Estes resultados indicam que, apesar de os participantes do grupo ativo terem efetivamente aprendido a realizar a tarefa desejada, não há evidências concretas da eficiência deste tipo de protocolo na aprendizagem de uma nova tarefa motora em indivíduos saudáveis.

Estudos de acompanhamento terão de ser realizados, de forma a comprovar as nossas suposições iniciais. A baixa amostra de sujeitos na realização deste estudo pode ter sido um fator de peso nos resultados obtidos, pois leva a uma maior influência da variabilidade dos dados nos mesmos. Para além disso, é necessário desenvolver técnicas de processamento de sinal mais eficazes pois, apesar de terem sido tomadas medidas para remover os artefatos de movimento e musculares do sinal adquirido, as figuras obtidas manifestavam corrupção por artefatos em canais periféricos. Finalmente, um refinamento do design da experiência pode permitir um melhor controlo de influências externas à tarefa e aumentar o nível de motivação e envolvimento dos participantes, permitindo, desta maneira, diferenciar as várias condições de forma mais acentuada.

Palavras-chave: Interfaces Cérebro-Máquina, Eletroencefalografia, Aprendizagem Motora, Feedback Robótico, Plasticidade Cerebral, Neuromodulação, Reabilitação, Acidente Vascular Cerebral
# Contents

Acknowledgements i  
Abstract iii  
Resumo vi  
List of Abbreviations ix  
List of Figures xi  
List of Tables xvii  

## 1 Introduction  

2 Background  

2.1 Stroke  

2.1.1 Clinical Aspects  
2.1.2 Rehabilitation Techniques  

2.2 Brain-Computer Interfaces  

2.2.1 General description and application  
2.2.2 Electroencephalography and its application to BCIs  

2.3 State of the art of EEG-based BCIs for motor recovery  

2.4 Project Aim & Motivation  

## 3 Methods  

3.1 Participants  

3.2 Data Acquisition  

3.2.1 Hardware  
3.2.2 Software  

3.3 Experimental Design  

3.4 Signal Processing & Analysis  

3.4.1 Online  
3.4.2 Offline  

## 4 Results and Discussion  

4.1 Classifier Performance  

4.1.1 Rest vs. Toe Abduction  
4.1.2 Toe Abduction vs. Toe Flexion  

vii
<table>
<thead>
<tr>
<th>4.2</th>
<th>Sensorimotor rhythms strength (ERD effect)</th>
<th>38</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.1</td>
<td>Time-frequency Analysis</td>
<td>38</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Frequency Analysis</td>
<td>41</td>
</tr>
<tr>
<td>4.3</td>
<td>Range of Motion (ROM)</td>
<td>46</td>
</tr>
<tr>
<td>4.4</td>
<td>EMG strength</td>
<td>50</td>
</tr>
</tbody>
</table>

5 Conclusion 57

Appendix 70
List of Abbreviations

ANOVA Analysis of Variance
AUC Area Under the Curve
BCI Brain-Computer Interface
BSS Blind Source Separation
CAR Common Average Reference
CCA Canonical Correlation Analysis
CIMT Constraint-induced Movement Therapy
CNS Central Nervous System
CR Classification Performance Rate
CVA Cerebrovascular Accident
DFT Discrete Fourier Transform
DV Decision Value
ECSS Ethics Committee of the Faculty of Social Sciences
EEG Electroencephalography
EMG Electromyography
EOG Electrooculography
ERD Event-related Desynchronization
ERP Event-related Potential
ERS Event-related Synchronization
ERSP Event-related Spectral Perturbation
FN False Negative
FP False Positive
FPR False Positive Rate
ICA Independent Component Analysis
ITI Inter-trial-interval
LDA Linear Discriminant Analysis
MCP Multiple Comparisons Problem
MI Motor Imagery
PSD Power Spectral Density
pSMA Prefrontal-supplementary Motor Area
rLLR Regularized Linear Logistic Regression
ROC Receiver Operating Characteristic
ROM Range of Motion
RTB Return to Baseline
SD Standard Deviation
SMA Supplementary Motor Area
SMR Sensorimotor Rhythms
SNR Signal-to-noise Ratio
SVM Support Vector Machine
SSVEP Steady-state Visual Evoked Potential
TFR Time-frequency Representation
TMS Transcranial Magnetic Stimulation
TN True Negative
TP True Positive
TPR True Positive Rate
TSG Technical Support Group
VR Virtual Reality
List of Figures

2.1 Schematic of the main stages of the BCI cycle. [Taken from van Gerven et al. 2009] 6

2.2 Left view of the human brain. Several areas are highlighted such as the frontal lobe, which is involved in executive functions such as behavioral control, planning or problem solving, and the parietal lobe that plays an important role in the integration of sensory information. The sensory signals from all modalities of sensation are received in the cerebral cortex immediately posterior to the central sulcus, which is called the somatosensory cortex. In its turn, the anterior portion (which constitutes the posterior half of the frontal lobe) is called the motor cortex and is devoted almost entirely to control of muscle contractions and body movements [47]. [Taken from Ferrez (2007)] 8

2.3 Representation of the spectral modulation in the motor cortical region and of the motor homunculus. (a) Electrocorticographic signals recorded during overt action and kinesthetic imagery of the same movement, for a single subject. Spectral changes in cortical surface potentials during hand and tongue movement and imagery can be observed in the 8-32Hz frequency range. Taken from Miller et. al. 64]. (b) Motor homunculus, representing somatotopic organization and control of the contralateral half of the body, from the toes (medially) to the face and tongue (laterally). The size of the cortical representation for each body part reflects the precision of motor control. 11

2.4 Power density analysis in healthy volunteers, taken from García-Cossio (2015) 81. (a) Grand average power density analysis over the Cz electrode, for active/passive walking conditions vs. baseline. (b) Topographic distribution of the ERD/S effect in the \( \mu \) (8–12 Hz), \( \beta \) (15-30Hz) and low \( \gamma \) (30-40Hz) bands. 13

2.5 Illustration of the BCI feedback cycle for this experiment. The first stage is EEG signal acquisition, which is processed online in order to extract the relevant features that will lead to the detection of SMR modulation, at a specific topographical location and frequency range. The characteristics that distinguish movement from non-movement will allow the classifier to produce predictions, which will be ultimately translated into physical outputs: visual feedback on a computer monitor and, for some participants, afferent feedback with a robotic device. 14

3.1 Electrode placement (a) Electrode positioning used for the EEG acquisition. The 64 channel layout is in accordance to the international 10-20 system. (b) Electrode positioning used for the EOG acquisition. (c) Electrode positioning used for the EMG acquisition. The electrodes were used to measure the electrical activity from three distinct muscles of the leg and foot region: 1 - Flexor hallucis brevis, 2 - Abductor hallucis and 3 - Flexor hallucis longus. The fourth electrode was placed on the ankle, and used as a reference. 18
3.2 Toe abductor device. (a) Schematic of the device. The green component represents the force meter, the red component represents the servomotor that pulls the toe to the defined position and, finally, the dark purple component is attached to the strap that is associated to the toe. (b) Example of the montage on a participant: 1 - force meter, 2 - servomotor, 3 - strap, 4 - arduino board. (c) Schematic of the Arduino Uno montage [Taken from www.arduino.cc]. The servo motor is connected to three pins - a digital pin, in order to receive commands from the software, a ground pin and the 5V power line.

3.3 Illustration of the ROM measurement. The black dot represents the first metatarsophalangeal joint, that was used as a reference point in order to guarantee consistency across all measurements. On the left, the assessment of the resting position is represented. The medial side of the right big toe and the second toe were marked on the sheet of paper, and the distance between both points, in degrees, was registered. The same procedure was applied for the assessment of the maximum abduction position, as represented on the right side of the figure. The difference between both measurements was registered as the ROM value for that session. This procedure was repeat in the beginning and end of every session, for all participants.

3.4 Schematic of a single block, for the calibration phase. The block begins with a 2s instruction, indicating which task will the subject have to perform for this specific block (abduct the big toe, flex all their toes or not move). Then, an audio cue sounds for 3s, corresponding to the ITI period, after which subjects were asked to perform the task for a period of 5s. The trial ends with a second audio cue of 0.7s and, afterwards, the ITI period starts again. Each block consists of 3 trials.

3.5 Schematic of a single trial, for the online feedback phase. The trial begins with a 3s audio cue, after which subjects were asked not to move. If the prediction value obtained by the classifier was above the confidence threshold defined by the ROC curve, the point counter would increase at a rate of 4Hz. This baseline period lasted 6s and then a second audio cue of 1.5s indicated that, from that point on, participants should try to abduct their toe. Similar to the baseline period, the point counter would also increase if the classifier was confident on the performance. For participants that were placed in the active or sham groups, additional physical feedback was given through the use of the toe abductor device. The apparatus would pull the toe with the same prediction rate. This part of the task lasted 12s. Then, the initial 3s audio cue appeared, functioning as a return to baseline (RTB) period, and a new trial would begin.

3.6 Per-class amplitude-spectra plot, obtained after preprocessing the calibration data from the 4th session of subject s9. Spectral differences between both classes (‘Toe Abduction’ and ‘Rest’) are more accentuated around 10Hz in midcentral channels, that correspond to the motor cortex. However, some channels also show peak amplitude difference around 25Hz, within the β frequency range.

3.7 Per-class area under the curve (AUC) plot, obtained after preprocessing the calibration data from the 4th session of subject s9. Spectral differences between both classes (‘Toe Abduction’ and ‘Rest’) are more accentuated around 10Hz and 25Hz in midcentral channels, that correspond to the motor cortex.
3.8 ROC curve (left) and histogram of DVs (right), computed from the calibration data of the 4th session of subject s9. The ROC curve displays the chosen threshold value of 0.682, i.e. the optimal threshold value for a 1% false alarm rate. By choosing this threshold value, it is possible to achieve a sensitivity of around 42%. The histogram of the DVs allows for the visualization of the data distribution of both classes. 'Rest' was defined as the positive class, and 'Toe Abduction' as the negative class. The red dotted lines mark the optimal threshold value, above and below 0. A perfect classification would lead to all DVs from the negative class to be below -0.682, and all values belonging to the positive class to be above 0.682.

3.9 ROC curves computed from two sets of DVs of the 4th session of subject s9. (a) Uses data from the beginning of the first trial of the session. The ROC curve displays the chosen threshold value of 0.690, i.e. the optimal threshold value for a 1% false alarm rate. The sensitivity at this threshold is of around 41%. (b) Uses data from the end of the last trial of the session. The ROC curve displays the chosen threshold value of 7.971. This value has a 1% false alarm rate, however the sensitivity is also around that percentage.

4.1 Boxplot displaying the subject variation of each group, in terms of difference in classification performance values, between the 5th and 1st sessions, for the 'Toe Abduction' vs. 'Rest' tasks.

4.2 Boxplot displaying the subject variation of each group, in terms of difference in classification performance values, between the 5th and 1st sessions, for the 'Toe Abduction' vs. 'Toe Flexion' tasks.

4.3 TFR plot, computed by averaging all trials acquired from the Active group (all subjects and sessions were grouped together). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between both conditions (rest and abduction), and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis.

4.4 TFR plot, computed by averaging all trials acquired from the Visual group (all subjects and sessions were grouped together). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between both conditions (rest and abduction), and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis.
4.5 TFR plot, computed by averaging all trials acquired from the Sham group (all subjects and sessions were grouped together). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between both conditions (rest and abduction), and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis.

4.6 Average ERD/ERS topoplots, between 15Hz and 30Hz, for all groups and sessions. Each row, from (a) to (e), represents the ERD/ERS data (in percentage) from session 1 to 5.

4.7 Boxplot displaying the subject variation of each group, in terms of difference in ROM values obtained before the experimental sessions. The values displayed in degrees correspond to the difference in ROM between the 5th and 1st sessions.

4.8 Boxplot displaying the subject variation of each group, in terms of difference in ROM values obtained after the experimental sessions. The values displayed in degrees correspond to the difference in ROM between the 5th and 1st sessions.

4.9 Average muscle activity recorded from all subjects belonging to the Active group after the 1st (left) and 5th (right) sessions. The normalized EMG amplitude was computed as multiples of baseline amplitude from the flexor hallucis longus, abductor hallucis, and flexor hallucis brevis muscles. The baseline period was defined as the 0-6s interval, and the red dotted line indicates the beginning of the movement period of the trial, which extends until the end of the time axis.

4.10 Average muscle activity recorded from all subjects belonging to the Visual group after the 1st (left) and 5th (right) sessions. The normalized EMG amplitude was computed as multiples of baseline amplitude from the flexor hallucis longus, abductor hallucis, and flexor hallucis brevis muscles. The baseline period was defined as the 0-6s interval, and the red dotted line indicates the beginning of the movement period of the trial, which extends until the end of the time axis.

4.11 Average muscle activity recorded from all subjects belonging to the Sham group after the 1st (left) and 5th (right) sessions. The normalized EMG amplitude was computed as multiples of baseline amplitude from the flexor hallucis longus, abductor hallucis, and flexor hallucis brevis muscles. The baseline period was defined as the 0-6s interval, and the red dotted line indicates the beginning of the movement period of the trial, which extends until the end of the time axis.

4.12 Boxplot displaying the subject variation of each group, in terms of difference in normalized EMG amplitude values obtained from the muscle flexor hallucis longus. The values displayed correspond to the median difference in EMG values, taken from the movement part of the trials (7.5-19.5s), between the 5th and 1st sessions.

4.13 Boxplot displaying the subject variation of each group, in terms of difference in normalized EMG amplitude values obtained from the muscle abductor hallucis. The values displayed correspond to the median difference in EMG values, taken from the movement part of the trials (7.5-19.5s), between the 5th and 1st sessions.
4.14 Boxplot displaying the subject variation of each group, in terms of difference in normalized EMG amplitude values obtained from the muscle flexor hallucis brevis. The values displayed correspond to the median difference in EMG values, taken from the movement part of the trials (7.5-19.5s), between the 5th and 1st sessions.

A.1 Illustration of the intrinsic muscles located within the right leg and foot, taken from Seeley et al. [117]. Muscle names are in bold font, and the referred ones are responsible for the flexion, extension, abduction, and adduction of the toes and foot. (a) Superficial muscles of the right foot, where we can find the abductor hallucis inserted in the base of the proximal phalanx of the great toe. This muscle is responsible for the abduction of the great toe. (b) Deep muscles of the right foot, including the flexor hallucis brevis inserted in the two tendons near the proximal phalanx of the great toe. This muscle allows for the flexion of the great toe. (c) Posterior view of the deep muscles of the right calf, where we can observe the flexor hallucis longus inserted in the distal phalanx of the great toe. This deep muscle of the posterior compartment plantar flexes and inverts the foot, as well as flexes the toes.

A.2 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s2 of the Active group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Rest’ classes are indicated through the colorbar.

A.3 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s9 of the Active group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Rest’ classes are indicated through the colorbar.

A.4 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s5 of the Visual group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Rest’ classes are indicated through the colorbar.

A.5 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s7 of the Visual group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Rest’ classes are indicated through the colorbar.

A.6 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s17 of the Sham group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Rest’ classes are indicated through the colorbar.

A.7 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s26 of the Sham group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Rest’ classes are indicated through the colorbar.

A.8 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s2 of the Active group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Toe Flexion’ classes are indicated through the colorbar.
A.9 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s9 of the Active group.

The magnitude of the spectral differences between 'Toe Abduction' and 'Toe Flexion' classes are indicated through the colorbar. 78

A.10 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s5 of the Visual group.

The magnitude of the spectral differences between 'Toe Abduction' and 'Toe Flexion' classes are indicated through the colorbar. 79

A.11 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s7 of the Visual group.

The magnitude of the spectral differences between 'Toe Abduction' and 'Toe Flexion' classes are indicated through the colorbar. 80

A.12 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s17 of the Sham group.

The magnitude of the spectral differences between 'Toe Abduction' and 'Toe Flexion' classes are indicated through the colorbar. 81

A.13 Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s26 of the Sham group.

The magnitude of the spectral differences between 'Toe Abduction' and 'Toe Flexion' classes are indicated through the colorbar. 82

A.14 TFR plots, with the averaged data for each session of the Active group (from first to last). On the left, the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between rest and abduction, and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis. 83

A.15 TFR plots, with the averaged data for each session of the Visual group (from first to last). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between rest and abduction, and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis. 84

A.16 TFR plots, with the averaged data for each session of the Sham group (from first to last). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between rest and abduction, and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis. 85
List of Tables

4.1 Classifier performance values (in percentage) for all subjects of the Active group, in each session. The classifier was trained to distinguish 'Rest' epochs from 'Toe Abduction' epochs based on frequency information in the EEG. ........................................... 34

4.2 Classifier performance values (in percentage) for all subjects of the Visual group, in each session. The classifier was trained to distinguish 'Rest' epochs from 'Toe Abduction' epochs based on frequency information in the EEG. ................................. 34

4.3 Classifier performance values (in percentage) for all subjects of the Sham group, in each session. The classifier was trained to distinguish 'Rest' epochs from 'Toe Abduction' epochs based on frequency information in the EEG. ................................. 34

4.4 Classifier performance values (in percentage) for all subjects of the Active group, in each session. The classifier was trained to distinguish 'Toe Abduction' epochs from 'Toe Flexion' epochs based on frequency information in the EEG. ................................. 36

4.5 Classifier performance values (in percentage) for all subjects of the Visual group, in each session. The classifier was trained to distinguish 'Toe Abduction' epochs from 'Toe Flexion' epochs based on frequency information in the EEG. ................................. 37

4.6 Classifier performance values (in percentage) for all subjects of the Sham group, in each session. The classifier was trained to distinguish 'Toe Abduction' epochs from 'Toe Flexion' epochs based on frequency information in the EEG. ................................. 37

4.7 Average ERD/ERS values (in percentage) for all subjects of the Active group, in each session. The values presented are a result of the average ERD over 20 mid-central electrodes, in the 15-30Hz frequency range. .................................................... 45

4.8 Average ERD/ERS values (in percentage) for all subjects of the Visual group, in each session. The values presented are a result of the average ERD over 20 mid-central electrodes, in the 15-30Hz frequency range. .................................................... 45

4.9 Average ERD/ERS values (in percentage) for all subjects of the Sham group, in each session. The values presented are a result of the average ERD over 20 mid-central electrodes, in the 15-30Hz frequency range. .................................................... 45

4.10 ROM scores, in degrees, of subjects that belong to the Active group, measured prior to (on the left) and post (on the right) experimental sessions. .................................................... 47

4.11 ROM scores, in degrees, of subjects that belong to the Visual group, measured prior to (on the left) and post (on the right) experimental sessions. .................................................... 47

4.12 ROM scores, in degrees, of subjects that belong to the Sham group, measured prior to (on the left) and post (on the right) experimental sessions. .................................................... 48
1. Introduction

Of the several virtues that are essential for humans to achieve a sense of purpose and general well-being, one of the most valued is the state of independence. This trait lets us have control of our own lives and actions, and when lost can lead to harmful effects on our self-esteem. The loss of independence can be defined at different levels, being the most severe the stage where there is a physical dependency on a third party in order to fulfill day-to-day activities. This is the case of the majority of people that suffer a stroke. Stroke is a common global health-care problem, and the leading cause of acquired adult disability in high-income countries [1]. Although most patients survive the initial injury, the setbacks include limitations in daily life activities, reduced participation in society and professional ventures, and overall loss of mobility. Motor impairment after stroke affects the control of the upper or lower extremities and, therefore, much of the focus of stroke rehabilitation is concentrated on the recovery of the impaired movement and associated functions [2].

Brain-Computer Interfaces (BCIs) are systems that collect physical correlates of brain activity, processing them in real-time to provide the user with feedback related to a specific process that is occurring in the brain. A recent and promising application of this field of technology concerns the motor rehabilitation of stroke patients. This process can be seen as a form of motor learning, which refers to a somewhat permanent change in motor behavior due to practice or experience [3]. This type of learning is mediated by neuroplasticity - the ability of the central nervous system (CNS) to undergo structural and functional changes in response to new experiences [4]. In this way, BCI training can be a useful rehabilitation tool by providing multisensory feedback that will modulate the brain activity, induce user-dependent plastic changes in the lesioned hemisphere and, thus, promote the restoration of motor function after stroke.

The project described throughout this thesis intended to study the effects of neuromodulation in the motor learning of the lower limb, within the context of stroke rehabilitation. As access to stroke patients is limited, we resorted to healthy participants to develop a new model for motor recovery after stroke, and used this model to validate the effectiveness of a BCI-triggered robotic training protocol. The goal was to aid participants in learning how to control a previously unused, yet functional, muscle - the abductor hallucis - with the use of real-time neurofeedback. For this we developed a lower limb motor training paradigm and a robotic apparatus to give physical feedback. Behavioral and electrophysiological data was gathered from all participants, and analyzed in both an online and offline setting.

The structure of this thesis is as follows: Chapter 2 provides a theoretical background on the standard and current rehabilitation techniques that are applied in the context of motor recovery after stroke. Thus, a thorough explanation, within this framework, of BCI paradigms that involve motor tasks and corresponding state of the art is presented. Chapter 3 describes the methods used for the development of the project - software and hardware design, signal processing and analysis, experimental procedure and general implementation details. This leads to Chapter 4 that displays and discusses the results obtained for the several performance measurements gathered in this experiment, followed by Chapter 5 that summarizes the conclusions drawn from this dissertation.
2. **Background**

2.1 **Stroke**

2.1.1 **Clinical Aspects**

Cerebrovascular accident (CVA), commonly designated as stroke, is one of the leading causes of mortality in most developed countries, responsible for more than 5.54 million deaths worldwide each year [5]. The occurrence of this and other types of cardiovascular diseases, such as hypertension and heart failure, is expected to increase substantially in the future [6]. However, the main burden of this condition is its debilitating character – stroke is the leading cause of chronic adult disability. It is estimated that 40% of stroke survivors are left with some degree of functional impairment [7]. The most common effect is motor impairment, which can be seen as a decrease in muscle control, leading to a loss of mobility.

There are two main types of stroke, hemorrhagic and ischemic, that can be caused due to intracerebral hemorrhage or from an obstruction of blood flow to the brain. Around 80% of occurring strokes are ischemic - the interruption of blood supply deprives neurons of oxygen and other necessary metabolic substrates which leads to cell necrosis and the presence of structural damage after only a few minutes [8, 9]. Depending on the location and extent of tissue damage, different neurological deficits can arise, causing impairments of sensory and motor functions such as visual-spatial disorders, cognitive deficits, spasticity or hemiparesis.

Even though stroke can cause devastating damage, many patients survive the initial event and undergo some spontaneous recovery. The extent of the post-stroke functional recovery is highly variable, as it is influenced by factors such as the age of the patient, the location and severity of the lesion, and anatomical or functional variations of their specific neural networks. In the first few months after the event, plastic changes take place in the brain, namely synaptic remodeling and dendritic sprouting. Post-stroke recovery depends on this ability of the CNS to modify its structure and function in response to external stimuli, i.e. experience. Such a dynamic reorganization of sensory and motor cortices is an important component in normal learning processes, as well as for recovery after neural injury. However, re-emergent post-stroke behavior is not identical to pre-stroke behavior, due to the loss of neurons with highly specific functions, but clinical data shows that functional recovery is possible when the neuroplasticity changes are further augmented by rehabilitative therapy [10].

2.1.2 **Rehabilitation Techniques**

The purpose of rehabilitation is to limit the impact of stroke-related brain damage on daily life, in order for survivors to reach the maximum physical and psychosocial recovery possible within the limits brought upon by their impairment. This can be achieved by using therapeutic and problem solving approaches that aim to maximize performance of daily activities and independence. Typically, early rehabilitation within the first few months relies on techniques that seek to promote changes in neuroplas-
ticity, whilst later rehabilitation encourages coping strategies based on educational and psychological theory [11].

Rehabilitation techniques, in a traditional sense, can be qualified as bottom-up approaches: the treatment focuses on the distal physical level, with the aim of influencing the neural system and its plasticity. Standard rehabilitation therapy relies on intensive repetition to achieve its goals. However, after completing it approximately 50-60% of stroke survivors experience some degree of motor impairment, and around 50% are at least partly dependent in daily activities [12]. These numbers show that current rehabilitation programmes have room for improvement.

Most of the recent protocols use the principles of motor learning, targeting deficits in the neuromuscular system and attempting to bridge them through practice. The term “motor learning” is used amongst different learning paradigms that comprise the (re)acquisition of a motor skill and aim to induce cortical reorganization. In general, these approaches encourage active patient involvement. Furthermore, they have been linked to the sprouting of dendrites, formation of new/alterations in existing synapses and production of neurochemicals. In this way, motor learning is associated with the type of neuroplastic changes that are thought to provide a mechanistic substrate to facilitate motor recovery after stroke [13]. If the practice method is repetitive, intensive and meaningful, patients demonstrate long-term retention of the re-acquired skill and the capacity to generalize the learned motor behavior [14].

Several movement therapy protocols have shown evidence of impact in motor and functional recovery.

Task-oriented therapy, also designated as task-specific training, is a term that originated from movement science and can be defined as a type of therapy where patients “practice context-specific motor tasks and receive some form of feedback” [15]. The emphasizes is on the training of functional tasks and not in the impairment (e.g. methods that focus on muscle strengthening). Previous animal studies demonstrated that training involving repetition of a specific motor movement, such as a skilled reaching task, can restore function by recruiting non-affected areas of the brain - adjacent to the lesion - or by resorting to supplementary brain regions [16]. Additionally, research regarding meaningful task-specific training has shown that this method can produce clinically relevant improvements when compared with standard training techniques [17]. These results suggest that the extent of performance increase is dependent on the specific external stimulus, and that generic use or activation does not induce the same degree of adaptive cortical reorganization of the CNS.

Constraint-induced movement therapy (CIMT) is another rehabilitation protocol that has shown to be efficient in promoting functional recovery, specifically in chronic stroke patients with paretic upper limb impairment [12]. Stroke patients tend to form the habit of not using the affected limb in daily activities. The CIMT method consists in constraining the unaffected limb, forcing individuals to practice tasks with the paretic arm. In this way, the therapy provides a repetitive practice of a task whilst increasing the number of successful efforts. However, the physiological functional alterations that are induced in the brain by CIMT, such as changes in cortical excitability or metabolic rate, fluctuate rapidly over time, which may question their reliability [18]. Furthermore, this rehabilitation protocol requires residual muscle strength and control, and a considerable percentage (> 50%) of stroke patients do not fulfill these criteria and, therefore, cannot benefit from the treatment [19].

Some studies show that mental practice with motor imagery (MI) can increase the effectiveness of usual occupational or physiotherapy [20, 21]. MI can be defined as a cognitive process in which the subject imagines that he is performing a specific motor action without actually performing any overt motor output [22]. Mental practice can modulate neural activity in regions of the brain similar to the ones activated during actual movements, such as the supplementary and premotor areas, the basal ganglia, the
cerebellum and the cingulate and parietal cortical areas of the brain. These regions have a prominent role in the preparation and execution of motor tasks [23]. Additionally, MI does not lead to a generalized arousal: research on the subject shows evidence of movement-specific activation patterns [24]. However, Ietswaart et al (2011) evaluated the therapeutic benefit of mental practice in sub-acute stroke patients with moderate motor weakness, and their results showed no significant differences in the measured outcomes, suggesting that mental practice of motor tasks isolated from physical practice is not beneficial for motor recovery [25].

Researchers and clinicians have obtained good outcomes from rehabilitation programs that resort to the use of robotic devices, in aiding impairments of the upper and lower limbs [26]. Robotic training has the potential to increase the intensity and effectiveness of therapy, in a quite affordable manner. These devices can deliver safe and precise exercises, with task-specific goals and quantifiable measurement of subject performance. However, some end-effector devices carry out a specific movement for the patient, in a bottom-up approach. This means that while motor function may be enhanced, electromechanical and robot-assisted training after stroke is not more likely to improve muscle strength [27]. If the device is set at a constant speed and pattern, users tend to rely on the robot to perform the movement and, thus, reduce their muscular activity [28]. User engagement in the task is crucial for a successful rehabilitation.

In order to promote such user engagement, virtual reality (VR) technology has been developed in the last decades, allowing the user to behave in challenging (yet safe) environments whilst maintaining experimental control over stimulus delivery and measurements [29]. VR offers, then, clinical assessment and rehabilitation options that are not available with traditional methods. Some studies have reported that post-stroke patients typically spend less than one hour per day in formal therapy, due to fatigue, lack of motivation and cooperation when attending therapy [30]. Virtual environments may provide interesting and engaging tasks that are more motivating than the formal repetitive design of standard therapy protocols. For locomotion, for example, the use of such technology gives the user a sense of forward movement - although their actual position in space is constant - and some results demonstrated moderate improvements in gait speed, increase in leg muscle activity, and even symmetry during gait for users with lower limb motor impairments [31]. VR is then a potentially powerful tool to generate a multidimensional, multisensory computer generated environment, which can be explored in real time.

Therefore, in order to promote motor learning, multisensory feedback should be given to the patient. Multisensory feedback is essential for primary and higher order cortical operations [32]. The environment that surrounds us is a multisensory one. Thus, the brain is equipped with the capacity to perform automatic and simultaneous integration and processing of sensory information. When robotic devices use this interactive feedback, there is an increase in training engagement and the sensorimotor loop that was disrupted after stroke can be re-established [33]. A recent and promising multisensory feedback approach for motor recovery in stroke patients is the use of electroencephalography (EEG) signals for BCIs coupled to orthotic devices [34] [35].

2.2 Brain-Computer Interfaces

2.2.1 General description and application

BCIs are systems that can bypass conventional channels of communication (e.g., muscles and speech) to provide direct communication and control between the human brain and physical devices by translating different patterns of brain activity into commands in real time [36]. Signal recordings of brain activity used by BCIs can be either invasive - requiring surgery to implant electrodes directly on or inside the
cortex - or non-invasive. Due to its low cost and practicality, non-invasive EEG recordings have been the most frequently used signals to develop BCI systems (Section 2.2.2).

As mentioned, the basic idea of a BCI is to translate subject-elicited patterns of brain activity into corresponding commands. A typical BCI system can be conceptually divided into signal acquisition, pre-processing, feature extraction, and classification (Figure 2.1). The cycle is closed by the user observing the output, leading to a judgment regarding the appropriateness of the system’s behavior [37].

**Figure 2.1:** Schematic of the main stages of the BCI cycle. [Taken from van Gerven et al. 2009] [37].

**Signal Acquisition** - EEG signals can be collected with electrodes that are placed on the surface of the scalp. Generally, the electrodes are placed according to the standard 10–20 international system, which describes head surface locations via relative distances between cranial landmarks, and conductive gel is applied to improve the signal-to-noise ratio (SNR) [38]. It should be mentioned that one of the most important parts of signal acquisition is the collection of good data. If the recording is showing excessive noise or if a given channel/trial experiences some obvious abnormality, it should be discarded immediately because it will give rise to a troublesome analysis - bad data leads to bad results and conclusions.

**Signal Processing** - Raw EEG data are high-dimensional time series that present a low SNR. Therefore, advanced signal processing has to be applied to reveal the relevant parts of the signal. This signal processing incorporates three stages:

- **Pre-processing** - The acquired signals are first pre-processed in order to remove artifacts such as power line noise and body movement. The simplest and most widely used method to remove artifacts is filtering (low-pass, high-pass, bandpass, and notch filtering), which is appropriate to remove frequency-specific noise. However, this method might filter useful components of EEG signals in the same frequency band as artifacts. Another approach is the use of complex computational methods, such as Independent Component Analysis (ICA), which divides a mixed signal into its statistically independent components [39]. Due to the volume conduction, raw EEG scalp potentials are known to have a poor spatial resolution. Spatial and temporal filters are extremely
useful in single-trial analysis in order to improve the SNR by enhancing the control signal and/or reducing noise [40]. Proper selection of the best spatial filtering method to use is determined by the location and extent of the control signal, as well as of the sources of EEG (or non-EEG) noise. The latter are more complicated to locate, seeing as they are highly variable both across and within subjects. Thus, there is need for methods that permit the optimization of spatiotemporal filters, such as statistical methods that analyze the mean, variance or standard deviation of EEG signals and use thresholding and correlation to achieve good results.

- **Feature extraction** - In order to improve the performance of the BCI system’s classifier, features that can significantly distinguish different classes are extracted. These features can be divided into two main categories: features in time domain (i.e., amplitudes, mean value, covariance, etc., of the event-related signal) and features in frequency domain (i.e., frequency power spectra of EEG signals). Time-frequency representations (TFRs) combine both temporal and spectral features, by describing how spectral power varies over time.

- **Classification** - The decoder used in the system will translate the extracted features from EEG signals into an output command. The output can be of a continuous or discrete nature, which leads to a regression or classification problem, respectively. Several prediction algorithms have been developed, such as the linear discriminant analysis (LDA) and support vector machines (SVM) [41]. Nevertheless, the classifier performance is strongly dependent on the number of extracted features and the amount of available training data. If the former is too high and the latter too low, a common situation when dealing with neuroimaging data, the classifier will be prone to overfitting and result in poor performance. Different approaches can be used to address this problem – employ feature selection methods to limit the number of features used, or use regularization techniques – and thus generalize the behavior of the decoder [42].

**Output** – This is the final component of the BCI cycle, and provides the user with observable feedback about the predicted intention. In the presence of an output device, the generated information can allow the user to control it (e.g., neural prosthetics or wheelchair) [43, 44]. There are several types of output modalities, such as auditory, motor commands, text or vibrotactile, depending on the type of experiment design used and its main goal.

Despite their promising nature, the robustness of existing BCI systems is not satisfactory due to the non-stationary nature of EEG signals. The recorded signals can undergo considerable changes between training and feedback mode as well as during feedback itself. This might be caused by task differences between training and feedback, variability of the recording caused by drying gel or micro movements of the electrodes, plasticity of the brain due to the gain of experience with the task, or even modulation of cognitive states like attention, motivation and vigilance [45]. Some of these conditions are natural properties of the human brain, and cannot be avoided. Taking this under consideration, a possible research direction could be the online adaptation of the classifier, and preliminary results have shown the feasibility and advantage of this method [46].

2.2.2 Electroencephalography and its application to BCIs

Electrical recordings from the surface of the brain have demonstrated that this organ presents continuous electrical activity. Both the intensity and the patterns of this electrical activity are determined by distinct
cognitive states such as sleep, wakefulness or even brain diseases (e.g., epilepsy or psychoses), which influence the level of excitation of different parts of the brain – see Figure 2.2.

Figure 2.2: Left view of the human brain. Several areas are highlighted such as the frontal lobe, which is involved in executive functions such as behavioral control, planning or problem solving, and the parietal lobe that plays an important role in the integration of sensory information. The sensory signals from all modalities of sensation are received in the cerebral cortex immediately posterior to the central sulcus, which is called the somatosensory cortex. In its turn, the anterior portion (which constitutes the posterior half of the frontal lobe) is called the motor cortex and is devoted almost entirely to control of muscle contractions and body movements. [Taken from Ferrez (2007)]

EEG is a non-invasive method to record electrical activity from multiple electrodes placed along the scalp. It measures mostly voltage fluctuations during synaptic excitation of the dendrites of pyramidal neurons in the cerebral cortex. The differences in potential are caused by the electrical dipoles between the body of the neuron and neural branches. Due to the capability to reflect both the normal and abnormal electrical activity of the brain, EEG has been found to be a very powerful tool in the fields of neurology, clinical neurophysiology, and BCI.

Biomedical signals such as EEG recordings are typically subject to noise by signals of non-cerebral origin, designated as artifacts. There are several types of artifacts with different origins, such as eye movement artifacts, caused by the large potential difference between the cornea and retina that occur with reflexive eye movements and are usually noticed in the frontal electrodes. Another common occurrence are electromyographic (EMG) artifacts, short duration potentials caused by muscle activity, that appears due to facial or body movement. There also exist non-physiologic artifacts, which arise from outside the body. The most common arise due to electrical interference of power lines and equipment. This background interference is likely to appear in all channels and recordings made in the same environment and with the same hardware. Nonetheless, as mentioned in Section 2.2.1 these factors can be taken into account and their effects minimized.

The rhythmic activity found in EEG recordings can be divided into bandwidths of several types, with different frequencies. Delta (δ) waves present frequencies of less than 4 Hz, and typically occur in very deep sleep state in adults and premature babies. They can be found in the brain’s frontal region in adults, and in the posterior region in children. The oscillations with frequencies between 4 and 8 Hz are designated as theta (θ). They normally occur in the parietal and temporal regions, and are prompted during emotional stress in some adults. Rhythmical waves that occur within the 8 to 13 Hz range
are defined as alpha (\(\alpha\)). Generally, they are recorded as sinusoidal waves with larger amplitudes over posterior regions, present in roughly 95% of healthy adults, especially during eyes-closed rest \[50\]. The \(\alpha\) rhythm can be attenuated or even abolished by visual attention or other sensory stimuli, mental alerting activities (e.g., mental arithmetic) or anxiety. With the same bandwidth as \(\alpha\), mu (\(\mu\)) rhythms are recorded from scalp electrodes over the sensorimotor cortex. The attenuation of the resting EEG \(\mu\) oscillations reflects the desynchronization of underlying cell assemblies, i.e., an increase of the load on cells that are directly related with motor function control. Thus, the desynchronization of \(\mu\) rhythm has been observed in adult subjects when they are executing motor commands - being it an active, passive or a reflex movement -, mentally visualizing the performance of the movement (MI) and even when observing the movement of a third party \[51, 52\]. Similarly to \(\mu\) oscillations, beta (\(\beta\)) rhythm suppression over central sensorimotor areas has also been linked to active motor control, such as active walking \[53\].

These rhythms occur at frequencies between 13 and 30 Hz, and can also be recorded in parietal and frontal regions, in busy, active concentration or anxious thinking states. Finally, for high frequencies of 30 to 60 Hz the oscillations are designated as gamma (\(\gamma\)). It is believed that \(\gamma\) rhythms may be involved in high-level cognitive processes, including perception, problem solving, fear, and consciousness \[54\]. Additionally, \(\gamma\) oscillations in the 30-40 Hz range have been observed in the EEG and MEG studies prior to, or during, voluntary finger, forearm or leg movements although they were not phase-locked to movement onset \[55, 56\].

Generally, brain activity is irregular and no specific pattern can be discerned in the EEG. However, in certain instances distinct patterns do appear, some of which are characteristic of specific cognitive states or processes. These features are designated as signatures, and can be divided into evoked and induced responses. Evoked responses are characterized by precise time and phase-locking to the stimulus onset. Therefore, they can be detected by averaging repeated single-trial responses, even in the cases where the amplitude of the signal is small. Induced responses, on the other hand, are not phase-locked. Instead, the power of the signal is time-locked to the stimulus. This means that the power of specific frequency bands has to be calculated prior to averaging across trials \[37\].

The extraction of these characteristic signatures from EEG recordings is the basis of any BCI system. Over the years, several types of event-related potential (ERP) paradigms have been used in research because of their promising applications.

Due to its consistency, the P300 is the most commonly used evoked response in BCI speller systems. It appears as a prominent positive deflection (usually with a parietal focus) of the ongoing EEG about 300 ms after the occurrence of an infrequent stimuli interspersed with a frequent one \[57\]. Another example is the steady-state visual evoked potential (SSVEP), that appears when stimuli modulated at a fixed frequency are presented in a rapid succession, and an increase in EEG activity arises at the stimulus frequency. One of the possible applications of SSVEP is the control of mobile robots \[58\].

Paradigms that involve induced responses use event-related desynchronization (ERD) and synchronization (ERS). These occur due to changes in the oscillatory behavior of groups of neurons, translating the decrease or increase in power at specific frequency bands. This phenomenon is particularly useful for the study of motor tasks, when measuring the ERD/S of \(\mu\) and \(\beta\) rhythms over the sensorimotor cortex \[51\].

### 2.3 State of the art of EEG-based BCIs for motor recovery

As previously mentioned in Section 2.1, stroke is one of the leading causes of disability in industrialized countries. Survivors can suffer several neurological deficits or impairments, which have a debilitating
impact in their daily life. Hemiplegia is the most common impairment after stroke, contributing significantly to the reduction in gait performance. In fact, many patients do not reach a walking level that enables them to perform regular daily activities [26].

BCI systems are a potentially powerful tool to incorporate as a top-down approach for neurorehabilitation, seeing as they can record and translate useful properties of brain activity related to the state of recovery of the patients. Top-down approaches are based on effective rehabilitation elements, such as active participation or error-drive-learning, providing a direct link with the neural circuitries activated during movement [59, 60]. There is increasing evidence, from studies that resort to transcranial magnetic stimulation (TMS) or EEG, showing that the strengthening of the signal modulated in the motor cortex and other descending input circuits is crucial for functional walking in humans [61, 62]. Gait, as well as other natural sequences of movements, is the result of very complex interactions. The neural information generated at the spinal cord is processed at the cerebral cortex, and filtered through the musculoskeletal system to produce and modulate the required movements for walking, in response to the environment. Thus, influencing only the distal physical level with the aim of altering the neural system is reducing the mechanisms underlying neural plasticity to a very simple and ineffective bottom-up model.

One of the most commonly used biomarkers for assessing EEG-based BCIs is the modulation of sensorimotor rhythms (SMR). Voluntary movements are carried out due to the activity of numerous distinct cortical circuits, such as the prefrontal-supplementary motor area (pSMA). These circuits converge into the primary motor cortex, that releases motor commands to muscles and spinal motor neurons [63]. Different regions of the motor cortex control different areas of the human body, and the distinct proportions of the brain dedicated to processing motor functions are represented in the cortical homunculus (Figure 2.3 (b)). Therefore, SMR reflects cortical activity originated in sensorimotor areas. As mentioned beforehand, the ability to modulate SMR can be quantified by the presence of a power decrease in $\mu$ and $\beta$ oscillations, time-locked to an event, designated as ERD. This phenomenon is considered to represent corticospinal excitability, and has been observed over the contralateral primary sensorimotor area during MI tasks, actual and attempted movements [64, 65]. Figure 2.3 (a) illustrates an example of the decrease in activation of cortical potentials, following overt action and kinesthetic imagery of hand and tongue movements.

Several studies have shown that mental practice can be effective in optimizing the execution of movements in athletes [66, 67]. Mental practice is a training method by which MI (the internal reproduction of a specific motor action without any overt motor output) is used with the intention of improving performance. In other words, it is the imagined rehearsal of a motor act with the specific intent of learning or improving that act [68]. MI strategies can be divided into kinaesthetic and visual. The former implies that the subject has the feeling that they are performing the act, with all the sensory consequences of that first-person perspective. During the latter, the subject sees himself or another performing the movement as from a distance, in a third-person perspective. Both viewpoints are associated with the activation of common neural networks in the SMA, precentral gyrus and precuneus, but results have shown that kinaesthetic MI is more effective in the modulation of corticomotor excitability [69].

Based on the effects of mental practice, some investigators have proposed its use in neurological rehabilitation [70].

Jacobson was a pioneer in investigating the role of mental activity in the behavior of the neuromuscular system [71]. He demonstrated that there was an increase in muscular activity when subjects were imagining movements. However, the magnitude of the activation was only a fraction of the activation that took place during actual performance.

Later studies, such as the ones conducted by Stippich et al. (2002) and Ehrssom et al. (2003),
Figure 2.3: Representation of the spectral modulation in the motor cortical region and of the motor homunculus. (a) Electrocorticographic signals recorded during overt action and kinesthetic imagery of the same movement, for a single subject. Spectral changes in cortical surface potentials during hand and tongue movement and imagery can be observed in the 8-32Hz frequency range. Taken from Miller et al. [64]. (b) Motor homunculus, representing somatotopic organization and control of the contralateral half of the body, from the toes (medially) to the face and tongue (laterally). The size of the cortical representation for each body part reflects the precision of motor control.

demonstrated that the motor imagery of different body parts (foot, hand and tongue) activated the precentral gyrus in a somatotopic manner. Their data suggested that the imagined body part is reflected in a somewhat direct way in the pattern of cortical activation: imagery of the finger movement activated the finger area of the cortical motor homunculus, imagery of toe movements activated the foot region of the posterior part of the contralateral SMA and primary motor cortex, and imagery of tongue movements activated the tongue region of the primary motor cortex [72, 24]. Thus, MI leads to central activation patterns that are movement-specific, instead of generalized muscular arousal.

Due to these findings, most BCI studies instruct healthy subjects to exclusively perform motor imagery tasks, regardless of their motor abilities. But research on the topic has demonstrated that brain patterns whilst imagining movements are less distinguishable from rest when compared with motor execution patterns [73]. Additionally, the task may feel less natural and intuitive, which makes it more difficult to perform. Therefore, letting a motor-impaired individual attempt instead of imagine a certain movement may result in higher performance rates.

Previous research has shown that by learning to control ipsilesional desynchronization of SMR, stroke patients could successfully control orthosis attached to the paralyzed limb [74, 75]. In these studies, the association of brain oscillations and grasping movements visibly prompted subject’s motor learning. Ramos-Murgualdy et al. (2012) developed an online BCI coupled with a robotic hand exoskeleton for flexing and extending the fingers [75]. The experiment used healthy participants, instructing them to perform five different tasks of closing/opening the hand: MI with and without direct control of the orthosis, passive and active movements of the device, and rest. Thus, in some conditions the control of the robotic hand was directly linked to brain activity, in others the movements where not coincident. Results were analyzed offline and performance was defined as the difference in power of the SMR during the execution of the motor task and rest. It was observed that proprioceptive feedback (feeling and seeing hand movements) improved BCI performance significantly, i.e., they demonstrated that the use of contingent positive proprioceptive feedback BCI enhanced SMR desynchronization during motor tasks. In a follow-up study (2013), they evaluated the efficacy of BCI training coupled to
regular physiotherapy in the improvement of forearm and upper arm function, in stroke patients with severe paresis [34]. His study showed that the combination of these two approaches induced functional improvements in chronic stroke patients’ motor function, even if the patients did not display evidence of residual finger movements.

Fewer studies have been carried out concerning lower limb rehabilitation, when compared with upper limb. The two types of recovery differ due to the underlying mechanisms of gait control [76]. Lower limb signal detection is more challenging as a result of the position of its motor representation within the primary motor cortex. Despite this, some studies have successfully detected ERD/ERS patterns during attempted and passive foot movements in healthy subjects and paraplegic patients suffering from a complete spinal cord injury [77]. These patients, with interrupted efferent and afferent pathways, showed a diffuse and broad distributed ERD/ERS pattern during attempted foot movements, contrasting with the focal beta ERD/ERS pattern of healthy subjects. This might indicate that different overlapping cortical networks can take place whilst attempting/imagining a motor task, some depending on the peripheral pathways, others depending on cortical networks.

Neuper and Pfurtscheller (2001) investigated induced beta oscillations in distinct sensorimotor areas, representing the hand and the foot [78]. In this study, subjects were asked to perform voluntary self-paced dorsiflexion of the foot or extension/flexion of the index finger. In separate blocks, they would receive electrical stimulation of the tibial and median nerves, that would allow them to execute the movement non-voluntarily. They reported that foot movements elicited a mid-central beta ERS with center frequencies of 21.5 ± 3.3 Hz, higher than the frequencies of 17.4 ± 1.8 Hz observed in hand movements. Beta oscillations induced by movement or stimulation of upper versus lower limb were best recorded over the corresponding cortical representation area, but they were also present in adjacent areas within the sensorimotor strip.

Locomotion is a major part of one’s independence. As previously mentioned, stroke patients with walking disorders might benefit from the use of BCI robot-assisted protocols. One of the most common impairments in stroke patients is the inability to perform dorsiflexion of the ankle during the swing phase of gait [79]. Studies have shown that robot training of ankle dorsiflexion improved patients paretic motor control and walking velocity, which indicates that targeting this type of motor control may be a key factor in locomotor therapies [80].

In order to evaluate the feasibility to decode walking intention from cortical patterns during robot-assisted gait training, García-Cossio et al. (2015) investigated the EEG spectral patterns related to robot-assisted active and passive walking in healthy volunteers and acute stroke patients [81]. Results showed that the control of the assistive gait device was feasible, with classification performances of 93% for healthy subjects and 89% for stroke patients, despite the presence of prominent muscle and movement artifacts during walking. Additionally, brain activity within the β frequency band detected from the Cz electrode was significantly different during passive and active walking when compared to baseline. Figure 2.4 shows this effect, as well as the topographical difference between the activity in the μ-band (more laterализed over the hand areas), and the β and γ effect (more medially focused over the foot regions). In this study, they also observed that the healthy volunteers’ modulation of low gamma activity in central midline areas was associated with the gait cycle phases, but not in the stroke patients. The hypothesis suggested is that the lack of low gamma modulation during the gait cycle might be related to the sensorimotor integration deficits presented in these patients.

Bearing all this in mind, this project (summarized in Section 2.4, and described more thoroughly throughout this thesis) intends to investigate how to enhance motor learning by resorting to an EEG-based BCI system coupled with a robotic device. Given that stroke patients struggle with reacquiring
lost motor skills, we aim to define a lower limb task that is extremely difficult to execute by healthy volunteers. Mulder et al. (2004), in order to gain more insight into the mechanisms underlying mental practice, proposed a task that incentivized healthy participants to learn a totally novel movement through mental imagery [68]. Subjects had to learn to perform the abduction of the big toe, without moving the other toes or the foot. Results showed that subjects that had no previous experience in executing the movement could not acquire the toe-abduction movement by means of mental practice. Only those who physically practiced the target movement improved significantly. When subjects had some experience in the task (could perform the movement to some extent), the performance rate grew significantly after mental practice as well as after physical practice.

2.4 Project Aim & Motivation

True motor recovery is achieved when alternate pathways are established, providing a new form of communication between the motor cortex and the distal muscles that were used before the injury.

This project intended to test the effectiveness of a BCI robot training protocol in the acquisition of a new motor skill. For stroke patients, recovery is a very strenuous and time-consuming process. Re-learning to perform movements that, prior to the disease, required no effort can be frustrating and difficult. Thus, the aim was to create a task that could, to some extent, recreate this type of learning obstacles in healthy individuals. The main goal was to aid healthy subjects in learning how to use functional muscles, and ultimately perform movements that they could not perform prior to training.
In the beginning of this project, pilot experiments were developed and carried out, in order to select the most appropriate novel motor skill to be used in a latter stage, and to guarantee the feasibility of detection of SMR modulation for the selected task. The chosen novel movement was toe abduction, task previously used in a study conducted by Mulder and colleagues, which showed encouraging results [68]. EEG signals were used to record SMR modulation, and EMG signals were used to observe the extent of motor unit activation. Ability to abduct the big toe was determined by measuring the range of motion (ROM) before and after the experimental training sessions. Figure 2.5 depicts the developed experimental neurofeedback cycle.

![Image of the experimental neurofeedback cycle](image)

Figure 2.5: Illustration of the BCI feedback cycle for this experiment. The first stage is EEG signal acquisition, which is processed online in order to extract the relevant features that will lead to the detection of SMR modulation, at a specific topographical location and frequency range. The characteristics that distinguish movement from non-movement will allow the classifier to produce predictions, which will be ultimately translated into physical outputs: visual feedback on a computer monitor and, for some participants, afferent feedback with a robotic device.

By providing brain signature based feedback with a robotic device, users were expected to acquire the new skill faster and more efficiently, in a top-down approach [26]. Recent studies have shown that proprioceptive afferent feedback can be a key component when attempting to close the sensorimotor loop and, thus, enhance rehabilitation effects [82]. Additionally, active participation in motor tasks has shown more efficient results in motor function gain, when compared to passive motor trainings [83]. This evidence indicates that the control of a robotic apparatus with active volition is one of the most promising approaches for motor neurorehabilitation.

In this way, the intent was to:

- Analyze changes in neural responses as well as behavioral responses of healthy volunteers;
- Establish the most suitable ERD signatures (best location and frequency range) for BCI control;
- Investigate how BCI-robot training can influence neuroplasticity during the acquisition of a new motor skill;
• Demonstrate that desynchronization of SMR is a reliable EEG signature to evaluate motor learning.

To achieve the abovementioned goals, EEG measurements were performed in 21 healthy participants, each taking part in 5 experimental sessions. Participants were equally divided into 3 groups, in order to assess the differences between continuous proprioceptive feedback, continuous visual feedback, and involuntary proprioceptive feedback. The developed system provided online feedback in real-time, but offline analysis of the acquired signals was also performed, in order to analyze the ERD/ERS effect in relevant frequency ranges and to implement statistical tests on the data.

This first approach can be a starting point to better understand the underlying processes of motor learning and BCI-robot control, and in a near future - out of the scope of this thesis - the acquired knowledge can be transferred to stroke patient motor rehabilitation.
3. Methods

3.1 Participants

Twenty-one healthy subjects participated in the experiment (five male and sixteen female, mean age 21.8 ± 2.8 years). Subjects had no history of neurological or psychiatric disorders. All the participants gave their informed consent before taking part in the experiment. Only one subject (s2) had prior experience with BCI measurements. Prior to the experiment, subjects were randomly assigned to one of three possible groups: visual feedback, active feedback or the sham/control group. Groups were evenly numbered (7 subjects per group), and all participants took part in 5 experimental sessions. The study was conducted at the Donders Institute for Brain, Cognition and Behaviour, and approved by the Ethics Committee of the Faculty of Social Sciences (ECSS) at the Radboud University, Nijmegen.

3.2 Data Acquisition

3.2.1 Hardware

EEG signals were recorded from 64 sintered Ag/AgCl active electrodes at a 2048Hz sampling rate, and downsampled by a factor of 8 - to reduce the amount of data and, thus, enhance computational speed -, using a Biosemi ActiveTwo AD-box amplifier (University of Amsterdam, The Netherlands). Impedance of the electrodes was kept below 50KΩ. The electrodes were placed in a cap, to ensure a correct placement along the scalp, according to the international 10-20 system (see Figure 3.1 (a)). Electrolyte gel was applied to guarantee good conductivity of the signal between the scalp and the electrodes.

In order to remove eye movement artifacts from the brain data, electrooculography (EOG) signals were recorded with Ag/AgCl surface electrodes. These were positioned around the eyes, as displayed in Figure 3.1 (b).

Muscle activity from the right leg and foot was recorded using three active surface electrodes, over the Flexor hallucis longus (calf), Flexor hallucis brevis (sole of the foot) and Abductor hallucis (lateral area of the foot), as seen in Figure 3.1 (c). Both the EMG and the EOG electrodes used were connected to the Biosemi ActiveTwo AD-box amplifier and, subsequently, had the same hardware and software specifications as the EEG electrodes.

In order to provide physical feedback of the desired motor task to the participants, a prototype of a toe abductor robotic device was developed. The apparatus executes the required movement of the target muscle of the right foot, the abductor hallucis, upon command in a systematic way. The device performs the movement with a limited amount of force and range of motion, as defined a priori in the software, to insure the safety of the user. The construction includes an Arduino Uno board, a servomotor, a force meter and a velcro strap for the toe (see Figure 3.2). As an additional safety measure, the user is able to remove their foot from the device at anytime. The device was designed and constructed in cooperation
with the Technical Support Group (TSG) of the Faculty of Social Sciences at the Radboud University, to ensure subject safety and was only used after their internal safety audit.

Figure 3.1: Electrode placement (a) Electrode positioning used for the EEG acquisition. The 64 channel layout is in accordance to the international 10-20 system. (b) Electrode positioning used for the EOG acquisition. (c) Electrode positioning used for the EMG acquisition. The electrodes were used to measure the electrical activity from three distinct muscles of the leg and foot region: 1 - Flexor hallucis brevis, 2 - Abductor hallucis and 3 - Flexor hallucis longus. The fourth electrode was placed on the ankle, and used as a reference.

Figure 3.2: Toe abductor device. (a) Schematic of the device. The green component represents the force meter, the red component represents the servomotor that pulls the toe to the defined position and, finally, the dark purple component is attached to the strap that is associated to the toe. (b) Example of the montage on a participant: 1 - force meter, 2 - servomotor, 3 - strap, 4 - arduino board. (c) Schematic of the Arduino Uno montage [Taken from www.arduino.cc]. The servo motor is connected to three pins - a digital pin, in order to receive commands from the software, a ground pin and the 5V power line.
3.2.2 Software

Experimental design and online control of the experiment were done in MATLAB 2015b and 2014b (MathWorks, Natick, MA). In order to implement the online neurofeedback BCI system, the Buffer BCI toolbox was used because of its functionality for real-time processing of neuroimaging data [https://github.com/jadref/buffer_bci](https://github.com/jadref/buffer_bci). The Buffer BCI is a platform independent framework, based on a client-server architecture with multiple clients obtaining and sending data to a central data and events server. This means that its functioning is dependent on three components: a running server (the buffer), the data acquisition client with the hardware to send data to the server, and the application client that will use the received data. The server is based on the Fieldtrip buffer specification for data access and storage [84].

Offline data analysis was done by resorting to both built-in and self-developed functions, as well as the Fieldtrip toolbox in the MATLAB 2015b development environment.

Finally, control of the toe abductor device was performed through an Arduino board, programmed to change the shaft position of the servomotor upon information received from the serial port. The information read by the input pin is a square pulse wave, updated every 20ms, with a pulse between 1 and 2 ms which corresponds to a change in position between 0 and 180°, respectively. MATLAB functions were developed to send position data to the Arduino hardware, allowing for the control of the rotation of the servomotor in real-time with high precision.

3.3 Experimental Design

Healthy participants attended five experimental sessions, within a 1-month time frame. Subjects sat in a comfortable chair, facing the computer monitor, and placed their right foot on the toe abductor device. Participants were randomly assigned to one of three possible groups, all using an EEG-based BCI system that relies on voluntary modulation of SMR when the user attempts to move the target muscle. The visual group received online visual feedback of their performance on the computer monitor. In the active group, successful modulation of SMR oscillatory activity had additional feedback, translated into the movement of the toe abductor apparatus. Finally, the sham group had the same montage as the active group, however the feedback was not coupled with the control of the robot. Subjects from the latter group were given mock feedback, i.e. a recording of the feedback given to another participants’ session. Because this was a novel movement that participants did not master, in the first session the movement was exemplified with the device, prior to the session, allowing for subjects to understand the correct way of performing this new motor skill.

In the beginning of the each session, after capfitting, participants were asked to abduct their toe without any external help, as much as they could, in order to measure the ROM, as illustrated in Figure [3.3]. The ROM was defined as the distance in degrees between adjacent toes. This distance was measured while the subjects were sitting on the chair, in an upright position, with their knees flexed approximately 120°. The right foot was placed on a sheet of millimeter paper, attached to the device. Subjects were instructed to press the lateral side of their right foot to a wooden board that is coupled to the device, in order to stabilize the foot. First, a straight line was traced along the first metatarsophalangeal joint, and used as a reference point. Afterwards, in order to measure the resting position, the medial side of the right big toe and the second toe were marked on the sheet of paper. Subsequently, subjects were asked to try to abduct the toe as far as possible, without moving the whole foot or the other toes. When maximum abduction was performed, the medial position of the first and second toes was marked again.
The difference, in degrees, between the resting position and abduction measurements was considered to be the ROM of that assessment. This protocol was based on the one used by Mulder et al. for ROM measurement [68].

Figure 3.3: Illustration of the ROM measurement. The black dot represents the first metatarsophalangeal joint, that was used as a reference point in order to guarantee consistency across all measurements. On the left, the assessment of the resting position is represented. The medial side of the right big toe and the second toe were marked on the sheet of paper, and the distance between both points, in degrees, was registered. The same procedure was applied for the assessment of the maximum abduction position, as represented on the right side of the figure. The difference between both measurements was registered as the ROM value for that session. This procedure was repeat in the beginning and end of every session, for all participants.

The experiment was divided into two parts: an initial calibration phase, and the main online feedback phase.

In the calibration phase, subjects were asked to perform three tasks on cue - flex all their toes, abduct the big toe or not move. This stage consisted in 18 blocks of 3 trials, each lasting 5s. The tasks were distributed equally throughout the blocks, and presented in a randomized order. In the beginning of the block, a visual display informed the participant of the motor task they should perform during the trials of that block. Participants were instructed to try to perform the movements as many times as they wanted to. Each trial start and ending was marked by an audio cue. Between trials, a 3s inter-trial-interval (ITI) was given, to avoid overlapping effects regarding signal modulation. During the trial, subjects looked at a fixation cross to prevent the occurrence of eye movement artifacts. Every 4 blocks, a pause period was given where the subject had to press a button in order to continue the experiment. Figure 3.4 illustrates the timeline of a single block of the calibration stage.

After completing the calibration, the data obtained from the abduction and resting tasks was saved and used to train a linear classifier. The data was also used to create a receiver operating characteristic (ROC) curve and, thus, calculate the adequate confidence threshold for the specific participant’s session. This threshold value was updated throughout the experiment. More details regarding this process can be found in Section 3.4.1.

The main stage of the experiment consisted of 30 blocks of 3 trials each. The timeline of a single trial is illustrated in Figure 3.5. During the blocks, online feedback of subject’s performance was given with a prediction rate of 4Hz. The type of feedback differed according to the group the participant
Figure 3.4: Schematic of a single block, for the calibration phase. The block begins with a 2s instruction, indicating which task will the subject have to perform for this specific block (abduct the big toe, flex all their toes or not move). Then, an audio cue sounds for 3s, corresponding to the ITI period, after which subjects were asked to perform the task for a period of 5s. The trial ends with a second audio cue of 0.7s and, afterwards, the ITI period starts again. Each block consists of 3 trials.

Figure 3.5: Schematic of a single trial, for the online feedback phase. The trial begins with a 3s audio cue, after which subjects were asked not to move. If the prediction value obtained by the classifier was above the confidence threshold defined by the ROC curve, the point counter would increase at a rate of 4Hz. This baseline period lasted 6s and then a second audio cue of 1.5s indicated that, from that point on, participants should try to abduct their toe. Similar to the baseline period, the point counter would also increase if the classifier was confident on the performance. For participants that were placed in the active or sham groups, additional physical feedback was given through the use of the toe abductor device. The apparatus would pull the toe with the same prediction rate. This part of the task lasted 12s. Then, the initial 3s audio cue appeared, functioning as a return to baseline (RTB) period, and a new trial would begin.
belonged to. For subjects that were in the visual group, feedback on performance was given in the form of points displayed on the screen. For the active and sham groups, besides the point counter on the screen, physical feedback was given with the toe abductor device. An initial example block was displayed in the beginning, allowing for users to practice the task beforehand and understand what they should do throughout the experiment. Each trial started with an audio cue of 3s, after which participants were asked to not move. This period of quiet baseline lasted 6s. Then, a second audio cue of 1.5s indicated that they should try to abduct their toe. Participants were instructed to try to perform the movement as many times as they wanted to, until they heard the final audio cue that indicated the end of the trial. This part of the trial lasted 12s. During the baseline and abducting periods a fixation cross was displayed on the screen. Resting periods were given between blocks, to avoid fatigue.

In the end of the experiment, the ROM was measured once more, to register any possible alterations in performance.

3.4 Signal Processing & Analysis

The acquired data was submitted to online and offline processing pipelines that, although similar, did differ in some aspects, due to the purpose of each process. The online processing steps were applied to guarantee that adequate feedback was given to participants, in real-time. In the offline analysis of the data, similar methods were applied, however the main goal was to obtain measurements that would indicate if significant changes in neural and behavioral responses were achieved.

3.4.1 Online

3.4.1.1 Calibration Data

In order to provide online feedback to participants, calibration data was acquired in the beginning of the experiment. As explained in Section 3.3, participants were asked to perform three different tasks on cue: flex all their toes, abduct the big toe or not move. Event labels were sent to the buffer, indicating the beginning and ending of each type of task. After completing all 18 blocks of calibration, a final "end" event label was sent, concluding this stage of the experiment.

A total of 36 trials (18 trials per task) of 5000 ms each, corresponding to the 'Toe Abduction' and 'Rest' tasks, were saved in a .mat file. These trials were subsequently cut into smaller 750 ms epoch windows, with an initial 500ms offset per trial. A new data structure, now containing 216 epochs, was saved and used to train a linear binary classifier.

3.4.1.2 Train Classifier

The classifier makes a classification decision based on the value obtained by the linear combination of the features of the training data. Due to the nature of the task at hand, a linear event-related spectral perturbation (ERSP) classifier was trained using the calibration data. ERSP classifiers are commonly used for motor imagery tasks, because they can distinguish features within the frequency domain, by identifying patterns, time-locked to the same events, of change in power over different frequency bins [85]. Additional details regarding the nature and functioning of the classifier used in the developed experiment will be described throughout this section.

Prior to training the classifier, the cut raw data obtained from calibration was preprocessed.

First, in order to remove linear trends from the data, the signal was detrended by subtracting the mean
value over time. This removes systematic shifts in the data and channel offsets, which might affect the analysis. Afterwards, to remove channels with excessive noise, electrode locations with a total power more than 3.5 standard deviations greater than the average channel power were identified as bad, and removed from the dataset. The remaining subset of channels were re-referenced by applying a Common Average Reference (CAR) spatial filter. Due to its computational simplicity, CAR is frequently applied to EEG data, allowing for the identification of small signal sources in very noisy recordings. This technique takes the average value of the entire electrode montage (the common average), and subtracts this value from that of the channel of interest. Seeing as the channels are equally spaced throughout the scalp, this method leads to a spatial voltage distribution with a mean of zero. In this way, it is possible to reduce the emphasis of components that are present in a large proportion of electrodes and, thus, functions as a high-pass spatial filter - i.e, accentuates components with highly focal distributions [40, 86]. A spatial whitening filter was also applied to the dataset, to remove any between-sensor correlations related to noise. Whitening is a linear transformation that converts the data - vector of random variables from which the covariance matrix is estimated - in such a way that, after whitening, it has an identity covariance matrix. Thus, after the transformation, the new whitened vector is composed by a set of variables that are uncorrelated (independent) and each have a variance of one [87, 88].

In an initial stage of this project, pilot experiments were carried out in order to understand and prevent possible issues that might affect the outcome of the experiment. One of the main issues that was noticed was the presence of EMG artifacts in the acquired signal. Due to the difficulty in performing this novel task, subjects tend to involuntarily contract their facial muscles, even when instructed to not do so. EMG signals have a broad frequency distribution, from 0 to more than 200Hz, nonetheless amplitude in peripheral electrodes on the scalp is greatest between 20 and 30Hz frontally and 40 to 80Hz temporally [89]. The fact that EMG spectra can present peaks within the $\beta$ frequency range can lead to unrecognized EMG contamination, mimicking actual EEG control and, consequently, entail misleading results. Thus, the detection of EMG artifacts requires appropriate spectral and topographical analyses. The EMG removal pipeline, that was applied to the dataset, was based on the findings of Clercq et al. [90]. Clercq and colleagues developed a novel method for muscle artifact removal in scalp EEG, demonstrating better results than the ones obtained after using low-pass filters or ICA - commonly used methods for EMG artifact suppression [91]. The method is based on statistical Canonical Correlation Analysis (CCA), applied as a blind source separation (BSS) technique [92, 93]. BSS-CCA relies on the fact that neural sources are mutually uncorrelated from each other, and highly temporally correlated with themselves, i.e. highly autocorrelated. Muscle artifacts, by comparison, have relatively low autocorrelation. Thus, the EMG artifacts can be extracted, and the relevant signal subsequently removed, by finding spatial filters that will minimize the correlation between the input data and its time-delayed version.

Afterwards, an EOG artifact removal pipeline was applied to the dataset. Ocular artifacts result from eye movements and blinks, demonstrating characteristics that are strongly non-stationary, and often localized with very large amplitude and low frequency. Additionally, its amplitude and duration can differ stochastically between successive eye movements/blinks. Hence, EOG artifact-contaminated EEG observations can show significant distribution changes in mean or covariance matrix [94]. Fortunately, EOG artifacts are highly localized (frontal electrodes) thus we can remove them from the data by 'regressing out’ the signal from the artifact-prone electrodes. For the regression function applied to the acquired signals, peripheral EEG channels were considered as potential artifact channels (‘AFz’, ‘AF3’, ‘FP1’, ‘FPz’, ‘FP2’, ‘AF4’, ‘AF8’, ‘AF7’, ‘Fz’, ‘O1’, ‘Oz’, ‘O2’, ‘P9’, ‘P10’), and the remaining electrode activity was decorrelated with-respect to these channels but ‘regressing out’ the artifact channels activity.

Finally, bad epochs were identified as those which had power more than 3 standard deviations (SD)
greater than the average epoch power. From the remaining subset of epochs, the power-spectral-density (PSD) was computed using Welch’s method. The time series data was divided in 250ms segments, and a periodogram spectrum of each segment was computed. The averaging of all the periodograms results in the Welch PSD estimate [95]. After obtaining this estimation of power, the data was sub-selected within the relevant frequency range for the task at hand (8-30Hz).

The processed data (750ms of clean EEG, collected from 64 channels, in 7 frequency bins between 8 and 30Hz, consisting of 64*7 = 360 features, on average) was then used to train a regularized linear logistic regression (rLLR) classifier, similar to the one used by Blokland and colleagues [96]. This type of classifier gives a natural estimate of the confidence in the prediction. The classifier uses a subset of data to find a linear weighting over the input features, by applying temporally stratified ten-fold cross-validation to set the regularization strength, allowing for the best fit to the data. For each fold, 90% of the trials were used for training the classifier, and 10% for testing it. The output of the classifier for each fold was the prediction decision values for each instance. In this way, it is possible to distinguish the specific pattern of spectral and spatial activation for the different tasks, and obtain accurate predictions.

In order to visualize the task related spectral changes, an average of the amplitude-spectra (Y) was calculated.

\[
Y(f, t) = \frac{1}{n} \sum_{k=1}^{n} (F_k(f, t))^2
\]  

(3.1)

Where \( n \) is the number of epochs, and \( F_k(f, t) \) is the spectral estimation of k-th epoch at frequency \( f \) and time \( t \). Mean amplitude-spectra values were calculated and displayed between 8 and 30Hz for 'Rest' and 'Toe Abduction' tasks. The result is a a two-dimensional image, illustrating the change in the frequency power spectrum between both classes at each data channel, as seen in Figure 3.6.

Figure 3.6: Per-class amplitude-spectra plot, obtained after preprocessing the calibration data from the 4th session of subject s9. Spectral differences between both classes ('Toe Abduction' and 'Rest') are more accentuated around 10Hz in midcentral channels, that correspond to the motor cortex. However, some channels also show peak amplitude difference around 25Hz, within the \( \beta \) frequency range.
Area under the ROC curve (AUC) plots were also computed, as represented in Figure 3.7. The AUC shows the frequency range where the difference between both classes is strongest (i.e. differs significantly from 0.5), allowing for the identification of the most relevant frequencies to differentiate 'Toe Abduction' from 'Rest'.

![Figure 3.7: Per-class area under the curve (AUC) plot, obtained after preprocessing the calibration data from the 4th session of subject s9. Spectral differences between both classes ('Toe Abduction' and 'Rest') are more accentuated around 10Hz and 25Hz in midcentral channels, that correspond to the motor cortex.](image)

The aforementioned figures represent the plots obtained after preprocessing the calibration data, for a single subject (s9) in one session. Thus, they are a visual representation of the average features that go into the classifier. The biggest difference between both classes can be found within motorstrip channels, around 10 and 25 Hz, as expected. In this case, the cross-validated classification performance was of 91.5%. The rLLR classifier’s decision values (DV) were calculated according to the following formula:

$$f(x) = \sum_{i=0}^{\infty} x(i)w(i) + b$$ (3.2)

Where the decision values $f(x)$ are computed by making a weighted summation of the input features $x$ and the classifier’s weighting $w$ over these features, with an offset $b$.

Additionally, the classification performance rate (CR) was measured by the balanced loss between both classes, which punishes stronger a wrong classification of an instance from the minority class than a wrong classification of an instance from the majority class, to prevent for unbalanced number of trials for each class. This is:

$$CR = \frac{TP + TN}{2}$$ (3.3)

This value falls in the range of 0 to 1, and after multiplying it by 100, to obtain a percentage value, we can state that an accuracy of 50% represented the chance level, i.e., no discrimination. Thus, the
classification performance takes into account the number of false positives (FP), false negatives (FN), true positives (TP), and true negatives (TN).

3.4.1.3 ROC Curve

The calibration data was used to produce a receiver operating characteristic (ROC) curve. ROC analysis is a statistical method for measuring performance in a binary classification task. It is an excellent visualization tool for understanding where the steepest trade-off exists between sensitivity and false alarms. Sensitivity, or true positive rate (TPR), measures the fraction of positive cases that are correctly identified as such. If, on the other hand, a fraction of negative cases are incorrectly identified as positive, this will count as part of the false positive rate (FPR). An ideal classifier will have a TPR of 1, whilst presenting a FPR of 0 [97, 98]. The ROC curve will then illustrate how the TPR/FPR vary, and allows us to select the most efficient threshold value for the task at hand.

Upon training the classifier, a results structure was saved with the DVs computed by the algorithm and the true labels for each epoch. This information was used to calculate the TPR and FPR for all values between the minimum and maximum DV from the results structure. The chosen threshold value was the value upon which the FPR was of 1%. This DV was then used as a reference to classify every new prediction in the beginning of the online neurofeedback stage, as explained in Subsection 3.4.1.4. Figure 3.8 illustrates the ROC curve and the comparative histogram for the calibration data of subject s9’s 4th session, and enables the comparison of the distributions of both classes ('Toe Abduction' - negative class - and 'Rest' - positive class).

![ROC curve and histogram](image)

Figure 3.8: ROC curve (left) and histogram of DVs (right), computed from the calibration data of the 4th session of subject s9. The ROC curve displays the chosen threshold value of 0.682, i.e. the optimal threshold value for a 1% false alarm rate. By choosing this threshold value, it is possible to achieve a sensitivity of around 42%. The histogram of the DVs allows for the visualization of the data distribution of both classes. 'Rest' was defined as the positive class, and 'Toe Abduction' as the negative class. The red dotted lines mark the optimal threshold value, above and below 0. A perfect classification would lead to all DVs from the negative class to be below -0.682, and all values belonging to the positive class to be above 0.682.

The ROC curve was computed at this stage and also throughout the online experiment, in order to define an appropriate classification threshold that adapted to the non-stationarity of the EEG, according to the data obtained for each subject in each session. Figure 3.9 displays two ROC curves, obtained
during the neurofeedback stage of subjects s9’s 4th session: (a) was computed with 226 DVs, obtained from the classifier in the beginning of the first trial. The optimal threshold found was of 0.690, value similar to the one obtained from the calibration data. (b) was computed with the same amount of DVs, from the end of the last trial, within the same session. The optimal threshold found was of 7.971, a value much higher than the ones presented in the beginning of the session. This means that the range of DVs obtained towards the end of the session was substantially bigger than the initial one. If we would have used the calibration data threshold value throughout the session, the number of misclassified samples would have been much higher.

Figure 3.9: ROC curves computed from two sets of DVs of the 4th session of subject s9. (a) Uses data from the beginning of the first trial of the session. The ROC curve displays the chosen threshold value of 0.690, i.e. the optimal threshold value for a 1% false alarm rate. The sensitivity at this threshold is of around 41%. (b) Uses data from the end of the last trial of the session. The ROC curve displays the chosen threshold value of 7.971. This value has a 1% false alarm rate, however the sensitivity is also around that percentage.

3.4.1.4 Real-time Feedback Data

After the aforementioned procedures were executed, the online continuous feedback stage began. The online processing pipeline was applied to 750ms moving data windows with a 500ms overlap - resulting in a 4Hz prediction rate - and used similar steps than the ones used to process the calibration data:

- Detrend;
- Spatial filtering (CAR + whitening);
- EMG artifact removal;
- EOG artifact removal;
- Bandpass filter (8-30 Hz);
- Welch window (250ms).

However, bad channels and trials were not removed online to avoid errors in the classifier’s adaptation to missing channels and data. Additionally, a bias adaptation filter was applied because this high-pass
filter removes slow drifts from inputs, by resorting to an exponential decay factor $\alpha$ for the moving average, as shown in the following equation:

$$f_x(t) = \frac{\sum_{i=0}^{\infty} (x(t-i) \times \alpha^i)}{\sum_{i=0}^{\infty} (\alpha^i)}$$  \hspace{1cm} (3.4)

Where $x(t)$ is the input data to filter and $\alpha$ the decay factor. The half-life value was set according to the duration of the experiment and the prediction rate, i.e. it corresponds to 3 minutes of data where the classifier is applied at a rate of 4Hz ($3 \times 60 \times 4$). The output of this function is the filtered data, $x(t) = x(t) - f_x(t)$.

Baseline drifts are a very common occurrence during EEG acquisition, caused by variations in temperature, bias in the instrumentation and amplifiers, or even because of changes in subjects’ alertness or cognitive engagement [49, 99]. Taking this into account, the ROC curve was continuously updated after every 10 predictions, by using the last 226 DV and true labels to calculate the TPR and FPR. In this way, it was possible to calculate the appropriate threshold to use for feedback throughout the experiment.

In order to give feedback, a logistic transformation was done online to the predicted DVs computed both by the classification algorithm and by the ROC curve. In this way, a more natural estimate of the class probability and the classifier’s confidence in the prediction was obtained. As we used a logistic regression classifier, this conversion from DVs ($f(x)$) to class probabilities ($Pr(x)$) was computed using the following logistic function:

$$Pr(x) = \frac{1}{1 + \exp(-f(x))}$$  \hspace{1cm} (3.5)

During the baseline period, if the predicted $Pr(x)$ was bigger or equal to the probability threshold obtained by the ROC curve, that would indicate that the classifier was confident in the positive class. Therefore, feedback was given in the form of points on the screen. On the other hand, during the movement period of the trial, if the predicted probability value was lower or equal to $1 -$ probability threshold, that would imply that the classifier was confident in the negative class. Hence, feedback was given again, in the form of points on the monitor and (for the active group) movements of the robotic device. This feedback system was not used for the sham group, since they received mock feedback (a playback of the feedback recorded from participants of the other two groups).

### 3.4.2 Offline

#### 3.4.2.1 EEG Analysis

The collected EEG data was analyzed per subject offline, in order to compare the signal modulation of the different groups, throughout the sessions.

Firstly, eye movement artifacts were removed from the trials by resorting to the signal recorded from the EOG electrodes placed around subject’s eyes. Signals correlated with the input signals from the data were discarded. Afterwards, linear trends were removed from the data (done per trial) and channels with a total power more than 3 SD greater than the average channel power were identified as bad and rejected from the dataset. The remaining channels were filtered between 8 and 40Hz, using a 6th-order bandpass butterworth filter, and re-referenced by applying a CAR spatial filter.

Finally, in order to remove any remaining artifacts, trials with an amplitude bigger than 50 $\mu$V were either rejected or, when possible, the channels that presented problematic trials were locally repaired. The repair of bad channels in the data was done by replacing them with a weighted average of the 3 to
8 nearest neighbor channels (the number of neighbor channels used for averaging was set according to the location of the target channel, as defined in the Fieldtrip template for the Biosemi 64-channel cap layout).

After these processing steps, the dataset was segmented based on the trigger codes into ‘Rest’ and 'Toe Abduction' trials, in order to compute the time and/or frequency plots presented in Chapter 4.

Event-related changes were computed and visualized by calculating TFRs of power. This was done using a sliding time window with a fixed length of 500 ms, which resulted in a frequency resolution of 2Hz. Thus, the power was calculated from 8 to 40Hz, for each frequency bin of 2Hz. Prior to calculating the power by discrete Fourier transformations the data was 'tapered'. This means that a hanning taper was multiplied to each time window of data, with the aim of reducing spectral leakage and control the frequency smoothing [100].

As mentioned, power estimates were calculated for multiple data segments, by using a sliding time window over the complete data segment, which resulted in a data structure with spatio-spectral-temporal nature. To reduce the dimensionality, the data was averaged over the timepoints that corresponded to each task (movement vs. non-movement) and the ERD/ERS values were computed. The ERD/ERS values were calculated in the same way as done by García-Cossio and colleagues (2015) - by normalizing the power in the frequency of interest, between 15-30Hz (β range), from the active movement condition by the corresponding baseline condition [81]. Because this relative measure is very sensitive to the outliers in the baseline period, the median power value was used to compute the ERD/ERS, since it is a more robust estimate than the mean. The following equation illustrates this procedure:

\[
ERD/ERS = \frac{\text{ToeAbduction} - \text{Rest}}{\text{Rest}} \times 100
\]

‘Toe Abduction’ corresponds to the moving segment of the trial, between 7.5 and 19.5s, whilst ‘Rest’ corresponds to the non-movement period of the trial, between 0 and 6s.

### 3.4.2.2 EMG Analysis

Muscle activity was recorded from three distinct muscles - *flexor hallucis longus, flexor hallucis brevis* and *abductor hallucis* - by using superficial electrodes for the measurements. The anatomic location of these muscles can be found in Figure A.1 of the Appendix. The collected EMG data was analyzed offline per subject and afterwards averaged per condition, in order to compare the signal amplitude of the different groups throughout the experimental sessions.

The processing of the muscle activity data began with the application of a Discrete Fourier Transform (DFT) filter, to remove the 50 Hz line noise and the harmonics at 100 and 150 Hz. Additionally, a 4th-order high-pass butterworth filter was applied with a cut-off frequency of 20Hz, to remove motion artifacts before subsequent processing. Following this, the signal measured from the ankle channel - used to provide a common reference/as a ground electrode and, thus, placed on electrically neutral tissue - was subtracted from the muscle electrodes. The data was then full-wave rectified by computing the absolute value of the raw signal. Generally, EMG signals fluctuate between positive and negative values during the movement of the muscle. Hence, the purpose of rectifying the signal is to ensure the raw signal does not average to zero. Full-length rectification adds the EMG signal below the baseline to the signal above it, making a conditioned signal that is all positive without losing significant information [101]. Finally, a 4th-order low-pass butterworth filter was applied to the data with a cut-off frequency of 20Hz, to temporally smooth the results after the rectification, and then averaged across trials for each sample and channel.
Grand average muscle activity was calculated first by normalizing the EMG signals across muscles for each participant and afterwards averaging per condition and session. For the normalization, the per-subject median baseline amplitude was calculated and this value was then used to divide the full trial signal. In this way, it was possible to obtain normalized plots (as seen in Section 4.4) corrected for subject-to-subject variations, where the EMG signal is reported in multiples of baseline amplitude.

### 3.4.2.3 Statistical Analysis

All reported data in Chapter 4 is presented per subject, however mean values ± SD are also indicated. Statistical evaluations of the different outcome measurements were done intra and intergroup-wise, by resorting to built-in and self-developed functions in the MATLAB 2015b and R development environments [102].

Within the same group, the first and last sessions were compared by using Wilcoxon signed-rank tests. Due to the low amount of subjects per group, we did not expect the populations to follow a normal distribution. The Wilcoxon test can be used as a non-parametric alternative of the paired sample t-test, allowing us to evaluate if two matched data samples (that come from repeated observations of the same subject) have an identical distribution. The null hypothesis is that the difference between the pairs follows a symmetric distribution around zero, and the alternative hypothesis is that this symmetric distribution does not occur [103]. This type of statistical analysis was used to compare both sessions per group in terms of classifier performance rates, ROM values, average ERD/ERS values within the beta frequency range (15-30Hz) for mid-central channels, and normalized EMG amplitude values of the target muscles.

Between groups, we assumed that the data samples were independent because they came from unrelated populations, i.e. the samples do not affect each other. Therefore, to compare the three conditions in terms of classifier performance rates, EMG and ROM values, the difference between the last and first sessions was calculated, and the resulting values were compared by resorting to a Kruskal-Wallis test. The Kruskal-Wallis test by rank is a non-parametric alternative to the one-way analysis of variance (ANOVA) test, allowing us to compare if two or more independent samples originate from the same distribution. A significant outcome from a Kruskal–Wallis test indicates that at least one sample stochastically dominates another sample, however the test does not identify where or for how many pairs does this stochastic dominance occur. Hence, the null hypothesis is that the medians of all groups are equal, and the alternative hypothesis is that at least one group presents a population median that is different from the population median of at least one other group [104]. In the cases where the Kruskal–Wallis test gave a significant outcome, a post-hoc analysis was performed to determine which levels of the independent variable differed from each other level. For this, a pairwise Mann–Whitney U-test, also known as a Wilcoxon rank-sum test, was applied to the data in order to compare the group levels whilst correcting for multiple testing. This non-parametric test can be used to determine whether two independent samples come from populations that present a same distribution - similar to the the Wilcoxon signed-rank test that is used on dependent samples.

A commonly occurring issue when performing statistical analysis of EEG data is the multiple comparisons problem (MCP). EEG data has a multidimensional structure, involving samples that originate from several timepoints and channels. Thus, the MCP arises from the fact that the effect of interest - difference between experimental conditions - will have to be evaluated between a large number of channel/time/frequency-pairs. This high number of statistical comparisons, when dealt with by resorting to standard statistical procedures, can lead to a high false alarm rate (i.e., the probability, under the hypothesis of no effect, of falsely concluding that there is a difference between the experimental conditions.

30
at one or more channel/time/frequency-pairs). A classically used approach to the MCP is to control the false alarm rate by setting a lower critical p-value ($\alpha$). This is called the Bonferroni correction. Hence, the critical value for an individual test is obtained by dividing the false alarm rate (usually 0.05) by the number of applied tests [105].

In order to compare the oscillatory brain activity between groups, for each session, cluster-based permutation tests were applied to the average ERD/ERS values of mid-central channels within the 15-30Hz frequency range. This allows us to identify differences in the topographic distributions and strength of the ERD effect across experimental conditions, within each session. This non-parametric test finds clusters of ERD/ERS strength where the topographic representation differs between conditions, while controlling for the false alarm rate due to the high numbers of comparisons. Cluster-based permutation tests were performed using Fieldtrip. The significant level $\alpha$ was 0.05 and we used 1000 permutations.
4. Results and Discussion

4.1 Classifier Performance

The calibration data collected from each subject, throughout all five sessions, was processed using the pipeline that is described in Section 3.4.1.2. Two different sets of binary classification performances were computed - first, as done in the online setting, the rLLR classifier was trained to distinguish the specific pattern of spatial and spectral activation for the 'Rest' and 'Toe Abduction' tasks. As an additional analysis, to observe if the differences between both movement classes differed for the duration of the sessions, the classifier was also trained to distinguish the 'Toe Abduction' and 'Toe Flexion' tasks.

4.1.1 Rest vs. Toe Abduction

Average classification accuracies (mean ± SD), for all groups and sessions, were above chance level (50%).

Table 4.1 shows the classifier performance values obtained for the subjects of the Active group, throughout the sessions. The average CR for the first session was of 78.9 ± 5.7%, similar to the average value for the fifth session (77.4 ± 4.2%). When comparing the values of both sessions, a p-value of 0.4982 (> 0.05) was obtained, indicating that no significant difference between the 1st and 5th sessions was found. The classifier results acquired from subjects belonging to the Visual group are represented in Table 4.2. Once again, the average CRs from the 1st and 5th sessions were not considerably different (77.4 ± 7.8% and 71.1 ± 4.7%, respectively). When applying a Wilcoxon signed rank test to compare both sessions, the result was a p-value of 0.2188 > 0.05, implying that there is no significant difference between sessions. Finally, the classification results from subjects of the Sham group are displayed in Table 4.3. The average CR of the 1st session was of 77.5 ± 9.0%, whilst the average CR of the 5th session was of 73.0 ± 5.8%. No significant difference between sessions was found for this group (p-value = 0.2188 > 0.05).

In order to compare the CR of the various conditions (Active, Visual and Sham), the difference between the classification performance values - distinguishing 'Toe Abduction' from 'Rest' - of the 5th and 1st sessions for each subject was computed. Figure 4.1 displays the distribution of this data in the form of boxplots, with the difference in CR of the three groups. The Active and Visual groups present the biggest subject variation, contrasting with the Sham group that has almost no disparity in the difference values. All median percentage values are negative, indicating that, on average, the CR of the 5th session was lower than the CR obtained in the 1st session, for all groups. The Active group presents the highest median CR values (around -3%). A Kruskal-Wallis test was applied to the computed values, however no significant difference between the mean classification accuracies of the three conditions was found ($\chi^2 = 1.0693$, p-value = 0.5859 > 0.05).
Table 4.1: Classifier performance values (in percentage) for all subjects of the Active group, in each session. The classifier was trained to distinguish 'Rest' epochs from 'Toe Abduction' epochs based on frequency information in the EEG.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>s2</td>
<td>70.8</td>
<td>77.5</td>
<td>74.2</td>
<td>77.0</td>
<td>76.3</td>
</tr>
<tr>
<td>s8</td>
<td>80.4</td>
<td>83.9</td>
<td>76.8</td>
<td>73.9</td>
<td>74.9</td>
</tr>
<tr>
<td>s9</td>
<td>76.3</td>
<td>82.7</td>
<td>81.4</td>
<td>91.5</td>
<td>83.0</td>
</tr>
<tr>
<td>s10</td>
<td>89.2</td>
<td>83.2</td>
<td>89.9</td>
<td>87.8</td>
<td>86.4</td>
</tr>
<tr>
<td>s13</td>
<td>86.9</td>
<td>87.8</td>
<td>71.0</td>
<td>79.9</td>
<td>77.8</td>
</tr>
<tr>
<td>s22</td>
<td>73.3</td>
<td>77.6</td>
<td>73.8</td>
<td>71.3</td>
<td>75.6</td>
</tr>
<tr>
<td>s24</td>
<td>75.2</td>
<td>82.0</td>
<td>78.3</td>
<td>73.2</td>
<td>68.1</td>
</tr>
</tbody>
</table>

Average $78.9 \pm 5.7$  $82.1 \pm 2.6$  $77.9 \pm 4.5$  $79.2 \pm 6.1$  $77.4 \pm 4.2$

Table 4.2: Classifier performance values (in percentage) for all subjects of the Visual group, in each session. The classifier was trained to distinguish 'Rest' epochs from 'Toe Abduction' epochs based on frequency information in the EEG.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>s5</td>
<td>77.1</td>
<td>74.9</td>
<td>70.0</td>
<td>79.3</td>
<td>79.8</td>
</tr>
<tr>
<td>s7</td>
<td>70.4</td>
<td>68.6</td>
<td>78.8</td>
<td>67.2</td>
<td>72.5</td>
</tr>
<tr>
<td>s11</td>
<td>84.1</td>
<td>83.0</td>
<td>78.0</td>
<td>70.7</td>
<td>74.2</td>
</tr>
<tr>
<td>s14</td>
<td>81.3</td>
<td>71.7</td>
<td>80.1</td>
<td>68.2</td>
<td>74.5</td>
</tr>
<tr>
<td>s15</td>
<td>93.8</td>
<td>68.7</td>
<td>78.4</td>
<td>78.2</td>
<td>69.5</td>
</tr>
<tr>
<td>s23</td>
<td>59.9</td>
<td>69.3</td>
<td>68.9</td>
<td>72.9</td>
<td>61.1</td>
</tr>
<tr>
<td>s28</td>
<td>74.9</td>
<td>66.8</td>
<td>65.8</td>
<td>72.0</td>
<td>66.4</td>
</tr>
</tbody>
</table>

Average $77.4 \pm 7.8$  $71.9 \pm 4.1$  $74.3 \pm 5.2$  $72.6 \pm 3.6$  $71.1 \pm 4.7$

Table 4.3: Classifier performance values (in percentage) for all subjects of the Sham group, in each session. The classifier was trained to distinguish 'Rest' epochs from 'Toe Abduction' epochs based on frequency information in the EEG.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>s17</td>
<td>68.8</td>
<td>66.5</td>
<td>67.4</td>
<td>70.1</td>
<td>64.6</td>
</tr>
<tr>
<td>s18</td>
<td>79.0</td>
<td>79.2</td>
<td>79.9</td>
<td>76.5</td>
<td>73.7</td>
</tr>
<tr>
<td>s19</td>
<td>58.2</td>
<td>63.7</td>
<td>70.0</td>
<td>75.1</td>
<td>64.7</td>
</tr>
<tr>
<td>s20</td>
<td>88.2</td>
<td>85.4</td>
<td>74.4</td>
<td>81.9</td>
<td>74.4</td>
</tr>
<tr>
<td>s21</td>
<td>79.5</td>
<td>80.3</td>
<td>83.9</td>
<td>77.5</td>
<td>74.6</td>
</tr>
<tr>
<td>s26</td>
<td>74.2</td>
<td>61.5</td>
<td>63.4</td>
<td>75.2</td>
<td>69.4</td>
</tr>
<tr>
<td>s27</td>
<td>94.8</td>
<td>94.7</td>
<td>95.4</td>
<td>91.6</td>
<td>89.9</td>
</tr>
</tbody>
</table>

Average $77.5 \pm 9.0$  $75.9 \pm 10.3$  $76.3 \pm 8.6$  $78.3 \pm 4.8$  $73.0 \pm 5.8$
Figure 4.1: Boxplot displaying the subject variation of each group, in terms of difference in classification performance values, between the 5th and 1st sessions, for the ‘Toe Abduction’ vs. ‘Rest’ tasks.

Overall, the classification performance values, differentiating ‘Toe Abduction’ from ‘Rest’, obtained throughout the five sessions for participants belonging to any of the groups, were above 70%. These values are in accordance to the CR values obtained in other BCI studies, showing the feasibility of binary classification, when attempting to distinguish lower limb movement tasks from baseline [106, 107].

Although the classification accuracies were above chance level, no significant differences were found between groups or within the same group, when comparing the first and last session. The average classifier performances for the three groups presented similar values, which might be related to the low amount of subjects within each group, increasing the effect of subject variability in the general group performance. The fact that the average CR values were not statistically different, between the first and last session of the same group, leads us to believe that the classifier is picking up information from broadly distributed signals, instead of specific frequency band or topographical location in the brain. However, if we observe the individual subject amplitude-spectra and AUC plots presented in Figures A.2 to A.7 of the Appendix, we can note that it is possible to discriminate both classes in terms of spectral characteristics, around 10Hz and 25Hz, for mid-central channels, and this efferent effect seems to become more visible - for some subjects - throughout sessions (e.g.: subject s9 in Figure A.5). On the other hand, other subjects show the opposite effect, i.e., the difference in spectral features between classes seems to became less noticeable in latter sessions (e.g.: subject s17 in Figure A.6). A possible explanation for the incoherency in both the classification outcomes and the visual inspection of the amplitude-spectra and AUC plots is the influence of afferent sensory information. EMG, movement and EOG artifacts could have affected the classification performance outcome, even though artifact removal functions were applied when training the classifier, to try to minimize these types of side effects.
4.1.2 Toe Abduction vs. Toe Flexion

Average binary classification accuracies (mean ± SD), for all groups and sessions, although still above the 50% chance level, were considerably lower than the values presented in Subsection 4.1.1.

The classifier results acquired from subjects belonging to the Active group are represented in Table 4.4. The average CRs from the 1st and 5th sessions were not considerably different (65.1 ± 3.4% and 66.2 ± 6.0%, respectively). When applying a Wilcoxon signed rank test to compare both sessions, the result was a p-value of 0.6875 > 0.05, implying that there is no significant difference between sessions. Table 4.5 shows the classifier performance values obtained for the subjects of the Visual group, throughout the sessions. The average CR for the first session was of 61.7 ± 4.4%, similar to the average value for the fifth session (63.0 ± 2.3%). When comparing the values of both sessions, a p-value of 0.4688 (> 0.05) was obtained, indicating that no significant difference between the 1st and 5th sessions was found. Lastly, the classification results from subjects of the Sham group are displayed in Table 4.6. The average CR of the 1st session was of 64.4 ± 5.7%, whilst the average CR of the 5th session was of 63.2 ± 5.2%. No significant difference between sessions was found for this group (p-value = 0.375 > 0.05).

With the purpose of comparing the CR of the distinct conditions, the difference between the classification performance values (distinguishing ‘Toe Abduction’ from ‘Toe Flexion’) of the 5th and 1st sessions of each subject was computed. Figure 4.2 shows the boxplots with the difference in CR of the three groups, exhibiting the distribution of this data. The Active group presents the biggest subject variation, with interquartile (from the 1st to the 3rd quartile) values ranging from around 7% to -4%. Contrary to what happened with the CR values comparing ‘Rest’ and ‘Toe Abduction’, the median percentage values of the Active and Visual groups were positive, indicating that, on average, the CR of the 5th session was higher than the CR obtained in the 1st session. In the case of the Sham group, the median value was negative (-2%). A Kruskal-Wallis test was applied to the computed values, however no significant difference between the median classification accuracies of the three conditions was found ($\chi^2 = 0.98701$, p-value = 0.6105 > 0.05).

Table 4.4: Classifier performance values (in percentage) for all subjects of the Active group, in each session. The classifier was trained to distinguish ‘Toe Abduction’ epochs from ‘Toe Flexion’ epochs based on frequency information in the EEG.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>s2</td>
<td>60.3</td>
<td>68.2</td>
<td>64.8</td>
<td>70.0</td>
<td>57.2</td>
</tr>
<tr>
<td>s8</td>
<td>69.2</td>
<td>76.2</td>
<td>70.3</td>
<td>69.8</td>
<td>63.7</td>
</tr>
<tr>
<td>s9</td>
<td>66.5</td>
<td>73.0</td>
<td>73.8</td>
<td>81.1</td>
<td>77.9</td>
</tr>
<tr>
<td>s10</td>
<td>61.7</td>
<td>50.0</td>
<td>62.8</td>
<td>68.1</td>
<td>63.9</td>
</tr>
<tr>
<td>s13</td>
<td>66.5</td>
<td>85.9</td>
<td>69.5</td>
<td>73.5</td>
<td>72.7</td>
</tr>
<tr>
<td>s22</td>
<td>61.3</td>
<td>61.6</td>
<td>67.0</td>
<td>59.6</td>
<td>68.8</td>
</tr>
<tr>
<td>s24</td>
<td>70.2</td>
<td>75.4</td>
<td>73.7</td>
<td>65.3</td>
<td>59.0</td>
</tr>
</tbody>
</table>

Average 65.1 ± 3.4 70.0 ± 8.7 68.5 ± 3.3 70.0 ± 4.9 66.2 ± 6.0

As mentioned beforehand, the classification performance values, differentiating ‘Toe Abduction’ from ‘Toe Flexion’, obtained throughout the five sessions for participants belonging to any of the groups, were considerably lower than the values acquired when comparing ‘Toe Abduction’ with ‘Rest’. This
Table 4.5: Classifier performance values (in percentage) for all subjects of the Visual group, in each session. The classifier was trained to distinguish 'Toe Abduction' epochs from 'Toe Flexion' epochs based on frequency information in the EEG.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>s5</td>
<td>65.4</td>
<td>71.0</td>
<td>58.1</td>
<td>65.5</td>
<td>68.0</td>
</tr>
<tr>
<td>s7</td>
<td>52.2</td>
<td>61.3</td>
<td>61.7</td>
<td>68.0</td>
<td>62.8</td>
</tr>
<tr>
<td>s11</td>
<td>63.7</td>
<td>58.5</td>
<td>65.4</td>
<td>69.0</td>
<td>61.3</td>
</tr>
<tr>
<td>s14</td>
<td>59.8</td>
<td>65.4</td>
<td>62.2</td>
<td>63.3</td>
<td>63.8</td>
</tr>
<tr>
<td>s15</td>
<td>67.1</td>
<td>63.3</td>
<td>55.6</td>
<td>57.1</td>
<td>57.2</td>
</tr>
<tr>
<td>s23</td>
<td>66.2</td>
<td>62.5</td>
<td>67.6</td>
<td>61.2</td>
<td>65.0</td>
</tr>
<tr>
<td>s28</td>
<td>57.8</td>
<td>55.8</td>
<td>53.7</td>
<td>59.5</td>
<td>62.6</td>
</tr>
</tbody>
</table>

Average 61.7 ± 4.4  62.5 ± 3.4  60.6 ± 4.1  63.4 ± 3.5  63.0 ± 2.3

Table 4.6: Classifier performance values (in percentage) for all subjects of the Sham group, in each session. The classifier was trained to distinguish 'Toe Abduction' epochs from 'Toe Flexion' epochs based on frequency information in the EEG.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>s17</td>
<td>56.0</td>
<td>58.9</td>
<td>64.8</td>
<td>69.1</td>
<td>58.7</td>
</tr>
<tr>
<td>s18</td>
<td>60.7</td>
<td>62.7</td>
<td>63.9</td>
<td>58.7</td>
<td>57.8</td>
</tr>
<tr>
<td>s19</td>
<td>56.7</td>
<td>63.2</td>
<td>NA</td>
<td>57.9</td>
<td>54.9</td>
</tr>
<tr>
<td>s20</td>
<td>74.8</td>
<td>84.7</td>
<td>74.8</td>
<td>69.3</td>
<td>76.3</td>
</tr>
<tr>
<td>s21</td>
<td>69.6</td>
<td>66.9</td>
<td>79.2</td>
<td>69.6</td>
<td>63.5</td>
</tr>
<tr>
<td>s26</td>
<td>64.3</td>
<td>61.3</td>
<td>60.0</td>
<td>67.2</td>
<td>67.1</td>
</tr>
<tr>
<td>s27</td>
<td>68.5</td>
<td>63.1</td>
<td>63.8</td>
<td>66.0</td>
<td>63.9</td>
</tr>
</tbody>
</table>

Average 64.4 ± 5.7  65.8 ± 5.7  67.8 ± 6.2  65.4 ± 4.1  63.2 ± 5.2

discrepancy in CR values between both binary classifications was expected, since it is very difficult to perform this novel motor task with the same precision as one would perform, for e.g., different finger movements. However, the fact that the classifier was able to distinguish both movements, even if not consistently throughout the whole population, shows that it is feasible (albeit more challenging) to detect differences in fine motor skills of the lower limbs.

No significant differences were found between groups or within the same group, when comparing the first and last session. Once again, the average classifier performances for the three groups presented similar values, that might be due to the low number of subjects per group which leads to a high subject variability within each group. The Active and Visual groups had an increase - even if statistically non-significant - in classification accuracy between the 1st and 5th sessions, that was not observed in the Sham group (which had a decrease in performance). This tendency in behavior possibly derives from the fact that these groups received online feedback that was linked to their brain activity, which might have incentivized participants to make a bigger effort when trying to perform the new motor skill correctly. The individual subject amplitude-spectra and AUC plots, presented in Figures A.8 to A.13 of the Appendix, show that for this type of binary classification problem, the discrimination between classes is less prominent than for the previous case. Some subjects present a difference in terms of spectral characteristics, around 10Hz and 25Hz, for mid-central channels (Figure A.9 and A.13), but most show no difference between classes within the motor cortex region or are corrupted by artifacts in more peripheral channels.
4.2 Sensorimotor rhythms strength (ERD effect)

The success of any BCI training paradigm is interlinked with the ability to recognize reliable EEG markers that constitute electrophysiological proxies for the relevant mental task discrimination. In order to obtain a clear distinction between classes, strong brain signatures, within consistent topographical locations in the head and frequency ranges, need to be identified. Thus, one of the primary outcome measures of brain plasticity obtained with this project was the ability to modulate SMR, characterized in terms of time-frequency representations and spectral information (i.e., ERD magnitude).

4.2.1 Time-frequency Analysis

Cortical activity accompanying voluntary movements is marked by the suppression of mu and beta (7–26 Hz) oscillations over the sensorimotor areas [63]. In order to visualize the event-related power changes, a normalization with respect to the baseline interval of each trial (0-6s) was performed. For each frequency, between 8 and 40Hz, the raw power values were expressed as the relative increase or decrease with respect to the power in the baseline interval (active movement period/baseline). Thus, the relative baseline is expressed as a ratio: values > 1 mean an increase in power, values < 1 indicate a decrease in power, and 1 represents no change between both periods. These power changes are displayed, per group, in the time-frequency plots of Figures 4.3, 4.4 and 4.5.
Figure 4.3: TFR plot, computed by averaging all trials acquired from the Active group (all subjects and sessions were grouped together). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between both conditions (rest and abduction), and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis.

Figure 4.4: TFR plot, computed by averaging all trials acquired from the Visual group (all subjects and sessions were grouped together). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between both conditions (rest and abduction), and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis.
Figure 4.5: TFR plot, computed by averaging all trials acquired from the Sham group (all subjects and sessions were grouped together). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between both conditions (rest and abduction), and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis.

The TFR for the Active group of Figure 4.3 shows that there is a clear desynchronization in the beta band during the movement period, when compared to the baseline period. This effect is more prominent in electrodes above the motor cortex and, as can be seen when we look at the single-channel (Cz) level, the strength of the effect is uniform throughout the whole movement time window. Additionally, we can observe a decrease in power in the alpha frequency band, between 8 and 13 Hz, within the sensorimotor region. However, this effect is less pronounced than the beta effect. Figure 4.4 represents the TRF for the Visual group. Similar to the previous condition, we can observe in the figure that there is a beta suppression after movement onset, within channels that belong to the motor cortex, although with less intensity and more evident a few seconds after the movement part of the trial starts. Lastly, Figure 4.5 shows the same type of plots for subjects belonging to the Sham group. Once again, this group presents a suppression of the beta band over sensorimotor areas, more evident in the first seconds after movement onset. However, as observed in the Visual group TFR plots, this effect (as well as the alpha suppression) is much less prominent than the effect verified in the Active group data.

The abovementioned figures are an average of all subjects and sessions for each group. As expected, the group that was subjected to an active neurofeedback paradigm, coupled to the robotic device, showed a clear ERD effect in the beta frequency range, over the motor cortex. We also hypothesized that this effect would occur for the subjects of the Visual group, which was the case, but with less intensity. The fact that the online feedback (for the Active group) was proprioceptive and visual during voluntary brain control, as opposed to the exclusive online visual feedback of the Visual condition, seems to have lead to a more effective closing of the loop between brain, movement and proprioception. However, the Sham condition, which did not receive feedback that was coupled to their brain activity, also produced a similar
effect. This might come from the fact that the afferent excitation of the sensorimotor brain through the robotic apparatus produces similar EEG frequency changes to the ones verified in previous conditions. In fact, previous studies have shown that passive movement affects frequency bands in a similar (yet somewhat weaker) way than active movement [108, 109].

Although measures were taken to exclude muscle and movement artifacts - which translated into the lack of channels in the plots, when compared to the initial 64 channel setting of the acquired data - it is still possible that some artifacts remain in the EEG data. The occipital and frontal channels show some increase in power for relatively low (<10 Hz) and high (>30 Hz) frequencies. Due to the locations within the head, it is not likely that these modulations are related to brain activity linked to the task at hand. These modulations could arise from muscle activity, or be caused by movement artifacts of the electrodes and head [110, 111].

Frequency bins from 15 to 30 Hz were used for subsequent analysis, based on the aforementioned observations indicating that the information in this frequency range most strongly reflects toe abduction brain activity in the subject population.

### 4.2.2 Frequency Analysis

The involvement of the cortex during the process of learning the novel motor skill was assessed by calculating ERD/ERS values, using the sensorimotor oscillations between 15 and 30 Hz (that correspond to the $\beta$ band). The ERD/ERS grand averages across participants over all the EEG electrodes, for all conditions and sessions are shown in the topoplots of Figure 4.6. The topographic pattern in all the conditions shows a desynchronization effect predominantly over centro-medial areas but this outcome is more prominent for subjects that belong to the Active group. For this condition, the pattern of brain activity seems to be scattered throughout the central channels in the 1$^{st}$ session (Figure 4.6 (a)), but by the 5$^{th}$ session the focus is split into two regions - one central and the other more lateralized towards the left hemisphere (Figure 4.6 (e)). With respect to the Visual group, the effect seems to evolve in the opposite way: in the 1$^{st}$ session it seems to show a tendency for a lateralization towards both hemispheres, but by the last session the effect seems to be more centralized. However this group, by visual inspection of the topoplots, seems to be the one with the less noticeable signal desynchronization. Finally, the ERD pattern for subjects belonging to the Sham group has a both a central and lateralized focus, and seems to oscillate in terms of signal strength throughout the sessions.

As mentioned in Subsection 3.4.2, cluster-based permutation tests were applied to the ERD/ERS datasets, in order to compare the different groups. For this comparison, only values from the 20 mid-central electrodes (FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2, C4, CP3, CP1, CPz, CP2, CP4, P3, P1, Pz, P2, P4) were used because this was the area that showed the strongest ERD effect in the computed topoplots, with the least corruption by scalp artifacts, and because it is the area where, according to previous literature, we would expect to see significant differences between conditions [81, 53]. The average of the values used for this comparison can be found in Tables 4.7 to 4.9. After using a 2-sided test with a significance level of 0.025 , on the datasets of the 1$^{st}$ session of all conditions, one negative cluster was found. However, this cluster was not significantly different (p-value = 0.1009). The same test was applied for the datasets of the 5$^{th}$ session, and this time four negative clusters were found. No significant difference was found between conditions for this session either (p-value = 0.0869, 0.3347, 0.3976, and 0.4545).

For the intragroup comparison, we also conducted cluster-based permutation tests to observe whether there was a systematic difference between the 1$^{st}$ and 5$^{th}$ sessions of participants subjected to the same
condition. We hypothesized that the 5th session would show a stronger desynchronization effect than the 1st session, thus a 1-sided (negative tailed) test was applied to the datasets, with a significance level of 0.05. For the Active group, four negative clusters were found. No significant differences were observed, however, between sessions (p-value = 0.4486, 0.5784, 0.8521, and 0.8791). The Visual group only presented one negative cluster and, once again, no significant difference was found (p-value = 0.5864).

Similar to what happened to the previous two conditions, no significant difference was found for the Sham group. This group presented six negative clusters (p-value = 0.5445, 0.5874, 0.8951, 0.9081, 0.9121, and 0.9261).

The fact that none of the performed comparisons were statistically significant was not expected, due to the visual inspection of the results obtained in the topoplots, the TRF plots of the previous section and of the Appendix (shown, per session and group, from Figure A.14 to A.16), and our own initial assumptions on the behavior of the populations. However, if we observe the values depicted on Tables 4.7 to 4.9, we can see that the difference in values is not as obvious as one would assume. In the first session, the Visual group presents the highest average value (i.e. weakest effect), followed by the Sham group and by the Active group, that shows the lowest average value (i.e. strongest beta suppression). Throughout the sessions, the Active condition has consistently the lowest values over the 20 channels that were analyzed. However, these are not that far apart from the average ERD/ERS values obtained from subjects belonging to the Sham condition. The Visual condition, by comparison, has the least prominent effect and, actually, increases its average value along all the sessions. Although the Active group presents the lowest values consistently all through the sessions, the SD is quite high. The higher the SD, the bigger the uncertainty of the goodness-of-fit of the data and, thus, less reliability can be put upon the results obtained in the experiment. This distribution of ERD/ERS values leads us to infer that there is a high variability amongst subjects belonging to the same group, which does not allow for a discrimination between groups. This issue could be tackled by gathering data from more subjects per group, in future experiments. Additionally, these values do not differ significantly between sessions contrarily to what has been published in previous neurofeedback studies that rely on MI to promote motor learning [112]. In these studies, when participants were subjected to several training sessions, significantly greater suppression of the SMR over task-related areas of the sensorimotor cortex was observed. This is an implicit assumption that permeates the neurofeedback literature: that the training process will lead to changes in the EEG, which in turn produces changes in behavior. However, as has been mentioned previously, the underlying complexity of neural dynamics can affect the outcome of such experiments. One factor that might have compromised our results was the lack of motivation of participants. When questioned afterwards, several participants stated that they felt that the task was tedious and repetitive, and that at times they got distracted or felt drowsy. In fact, some studies have shown the importance of motivation for the success of BCI applications [113, 114]. If subjects are engaged in the task at hand, they will be more focused on their performance and ultimately be able to modulate their brain activity in a more efficient way (e.g.: fMRI experiments showed that motivation/mental effort during a MI task can enhance the BOLD signal) [115].

A second interesting finding regarding the intergroup ERD/ERS effect is that we observed a different spatial distribution between the Active condition, when compared with the other two conditions. The lateralization towards the left hemisphere for subjects belonging to the Active group became more apparent throughout sessions. This lateral shift in beta suppression could be justified by the fact that participants, when trying to abduct their toe - in which they had no prior experience -, tended to slightly move their right leg, although they were instructed not to do so in the beginning of the experiment. This involuntary contraction/movement of the leg, whilst trying to achieve the best strategy to perform the novel move-
ment, might be more evident in participants from this group because it is the only condition where they were subjected to the mechanical pull of the toe (stronger efferent response than the Visual condition), whilst receiving brain-coupled feedback (contrarily to the Sham group). Some of the topoplots from the Sham condition also seem to indicate a tendency for lateralization of the beta suppression, which means that this might be an effect caused/emphasized by the design of the physical feedback device, but overall this observation is only consistent for the Active group data. Due to the non-significant results obtained, whether this difference is caused by inter-individual differences or represents a true difference in the intergroup spatial distribution of beta ERD remains to be investigated.
Figure 4.6: Average ERD/ERS topoplots, between 15Hz and 30Hz, for all groups and sessions. Each row, from (a) to (e), represents the ERD/ERS data (in percentage) from session 1 to 5.
Table 4.7: Average ERD/ERS values (in percentage) for all subjects of the Active group, in each session. The values presented are a result of the average ERD over 20 mid-central electrodes, in the 15-30Hz frequency range.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>s2</td>
<td>-20.53</td>
<td>-18.47</td>
<td>-19.03</td>
<td>-25.21</td>
<td>-23.07</td>
</tr>
<tr>
<td>s8</td>
<td>-15.86</td>
<td>-19.47</td>
<td>-18.08</td>
<td>-17.62</td>
<td>-13.84</td>
</tr>
<tr>
<td>s22</td>
<td>-8.09</td>
<td>1.17</td>
<td>-7.52</td>
<td>-9.24</td>
<td>-9.45</td>
</tr>
<tr>
<td>Average</td>
<td>-15.61 ± 3.91</td>
<td>-14.45 ± 5.97</td>
<td>-16.71 ± 5.50</td>
<td>-17.77 ± 5.80</td>
<td>-14.98 ± 8.16</td>
</tr>
</tbody>
</table>

Table 4.8: Average ERD/ERS values (in percentage) for all subjects of the Visual group, in each session. The values presented are a result of the average ERD over 20 mid-central electrodes, in the 15-30Hz frequency range.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>s5</td>
<td>-17.48</td>
<td>-24.66</td>
<td>-23.34</td>
<td>-28.66</td>
<td>-25.29</td>
</tr>
<tr>
<td>s7</td>
<td>2.73</td>
<td>-7.81</td>
<td>-1.92</td>
<td>2.79</td>
<td>-13.61</td>
</tr>
<tr>
<td>s11</td>
<td>-3.94</td>
<td>-3.94</td>
<td>2.25</td>
<td>-5.84</td>
<td>4.02</td>
</tr>
<tr>
<td>s14</td>
<td>-15.66</td>
<td>-9.09</td>
<td>-8.83</td>
<td>5.13</td>
<td>-3.32</td>
</tr>
<tr>
<td>s15</td>
<td>-27.46</td>
<td>-8.55</td>
<td>-7.44</td>
<td>-7.01</td>
<td>-9.02</td>
</tr>
<tr>
<td>s23</td>
<td>-5.84</td>
<td>2.05</td>
<td>-0.54</td>
<td>0.84</td>
<td>8.05</td>
</tr>
<tr>
<td>s28</td>
<td>3.06</td>
<td>-3.34</td>
<td>-11.06</td>
<td>-5.39</td>
<td>0.14</td>
</tr>
<tr>
<td>Average</td>
<td>-9.23 ± 9.41</td>
<td>-7.90 ± 5.31</td>
<td>-7.27 ± 6.17</td>
<td>-5.45 ± 7.19</td>
<td>-5.58 ± 8.91</td>
</tr>
</tbody>
</table>

Table 4.9: Average ERD/ERS values (in percentage) for all subjects of the Sham group, in each session. The values presented are a result of the average ERD over 20 mid-central electrodes, in the 15-30Hz frequency range.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>s17</td>
<td>-3.49</td>
<td>-7.15</td>
<td>-2.71</td>
<td>-4.40</td>
<td>-3.74</td>
</tr>
<tr>
<td>s18</td>
<td>-9.96</td>
<td>-8.70</td>
<td>2.37</td>
<td>-19.22</td>
<td>-11.40</td>
</tr>
<tr>
<td>s19</td>
<td>-8.89</td>
<td>-5.02</td>
<td>-4.00</td>
<td>-6.55</td>
<td>-6.52</td>
</tr>
<tr>
<td>s20</td>
<td>-6.07</td>
<td>-17.85</td>
<td>-8.06</td>
<td>-9.37</td>
<td>-11.59</td>
</tr>
<tr>
<td>s21</td>
<td>-15.07</td>
<td>-15.09</td>
<td>-17.43</td>
<td>-22.24</td>
<td>-21.82</td>
</tr>
<tr>
<td>s26</td>
<td>-12.08</td>
<td>-9.96</td>
<td>-6.56</td>
<td>-15.14</td>
<td>-18.04</td>
</tr>
<tr>
<td>s27</td>
<td>-25.53</td>
<td>-21.07</td>
<td>-27.57</td>
<td>-24.05</td>
<td>-18.48</td>
</tr>
</tbody>
</table>
4.3 Range of Motion (ROM)

With the purpose of obtaining a behavioral measurement of the performance of the novel motor task, ROM values related to the abduction of the big toe were registered per subject, prior to and post-session, based on the protocol used by Mulder and colleagues [68].

The ROM measured before and after each experimental session, for subjects belonging to the Active group, are presented in Table 4.10. A Wilcoxon signed-rank test was applied in order to compare the average angles between the 1st and 5th sessions of the group. A significant difference was found for values recorded prior to the sessions (p-value = 0.0355 < 0.05). However, when comparing both sessions in terms of post-measurement values, no significant difference was found (p-value = 0.1501 > 0.05).

Table 4.11 shows the ROM angles acquired from subjects of the Visual group. No significant differences were found between the first and last session, for values obtained before or after each experimental session (p-value = 0.0917 > 0.05 and p-value = 0.2463 > 0.05, respectively). Lastly, the angles collected from subjects that were part of the Sham group are reunited in Table 4.12. Once again, neither the prior (p-value = 0.4461 > 0.05) nor the post (p-value = 0.4004 > 0.05) ROM values were significantly different between sessions.

Similar to what was done with the CR, the difference in ROM values between the 5th and 1st sessions was computed. A Kruskal-Wallis test was applied to this data, in order to perform the comparison between the 3 groups. No significant difference was found between conditions for values gathered before the sessions ($\chi^2 = 3.6485$ , p-value = 0.1613 > 0.05) nor for values collected after the sessions ($\chi^2 = 0.4084$ , p-value = 0.8153 > 0.05). Figures 4.7 and 4.8 display the distribution of this difference for each group, regarding the prior and post ROM values, in the form of boxplots. In the former figure, we can observe that the Active and Visual groups have the biggest subject variability, and present the same mean difference value (which means that, on average, both populations had an increase of $5^\circ$ from the first to the last session). Both groups display skewed data, with the Visual condition having a wider range in values in the 3rd quartile - i.e., the data is more spread out for lower angle values - and the Active condition showing a wider range in the 1st quartile (in higher angle values). Although the Sham group displays a median of around $2^\circ$, part of the subjects show no difference in ROM between sessions, or even a decrease between the first and last measurement. This also happened for some subjects of the Visual group. In the latter figure, we can observe that the Sham condition displays the biggest subject variability, with a much bigger interquartile range (between $-2^\circ$ and $9^\circ$) than the one presented for the datasets of the other conditions. Once again the Sham group presents the lowest median ($0^\circ$), followed by the Visual group ($1^\circ$) and finally by the Active group ($2^\circ$).

The main target of the training sessions was to increase the ROM. As expected, the values measured for subjects that belonged to the Active group were, overall, higher than the values measured from participants that were subjected to the other two conditions. This was the only group that displayed a statistically significant increase between the first and last session. This result is in accordance with the results published by Mulder and colleagues (in 2004) [68]. They observed that, if subjects were completely unable to abduct their toe prior to the experiment, only those who physically practiced the target movement improved significantly. But, if - even to a small extent - participants could abduct their toe prior to engaging in the study, then either mental or physical practice were beneficial for the significant increase in ROM values. This lead them to conclude that the effects of mental practice leads to improvement only when a central representation of the target movement is present.
Table 4.10: ROM scores, in degrees, of subjects that belong to the Active group, measured prior to (on the left) and post (on the right) experimental sessions.

<table>
<thead>
<tr>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA</td>
<td>3</td>
<td>10</td>
<td>11</td>
<td>9</td>
<td>s2</td>
<td>4</td>
<td>7</td>
<td>11</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>13</td>
<td>2</td>
<td>5</td>
<td>s8</td>
<td>7</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>9</td>
<td>19</td>
<td>18</td>
<td>s9</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>15</td>
<td>10</td>
<td>13</td>
<td>s10</td>
<td>12</td>
<td>16</td>
<td>10</td>
<td>22</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>12</td>
<td>17</td>
<td>16</td>
<td>s13</td>
<td>8</td>
<td>14</td>
<td>17</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
<td>9</td>
<td>12</td>
<td>13.5</td>
<td>s22</td>
<td>7</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>3.5</td>
<td>8.5</td>
<td>1</td>
<td>7</td>
<td>s24</td>
<td>8</td>
<td>6</td>
<td>10</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.5±2.8</td>
<td>8.5±4.1</td>
<td>10.9±2.1</td>
<td>10.3±5.1</td>
<td>11.6±4.0</td>
</tr>
</tbody>
</table>

Table 4.11: ROM scores, in degrees, of subjects that belong to the Visual group, measured prior to (on the left) and post (on the right) experimental sessions.

<table>
<thead>
<tr>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
<th>Subjects</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
<th>Session 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>6</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>s5</td>
<td>6</td>
<td>2</td>
<td>8</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>9</td>
<td>15</td>
<td>8</td>
<td>s7</td>
<td>19</td>
<td>9</td>
<td>16</td>
<td>20</td>
<td>23</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>10</td>
<td>15</td>
<td>16</td>
<td>s11</td>
<td>14</td>
<td>21</td>
<td>15</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>10</td>
<td>s14</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>17</td>
<td>15</td>
<td>12</td>
<td>s15</td>
<td>11</td>
<td>15</td>
<td>10</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>8.5</td>
<td>8</td>
<td>5.5</td>
<td>7.5</td>
<td>s23</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>3</td>
<td>-1.5</td>
<td>2</td>
<td>-0.5</td>
<td>s28</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>-1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.5±2.1</td>
<td>7.4±3.0</td>
<td>7.2±4.3</td>
<td>9.1±5.1</td>
<td>8.7±3.4</td>
</tr>
</tbody>
</table>
Table 4.12: ROM scores, in degrees, of subjects that belong to the Sham group, measured prior to (on the left) and post (on the right) experimental sessions.

<table>
<thead>
<tr>
<th></th>
<th>Prior ROM values (°)</th>
<th>Subjects</th>
<th>Post ROM values (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Session 1</td>
<td>Session 2</td>
<td>Session 3</td>
</tr>
<tr>
<td>15</td>
<td>11</td>
<td>7.5</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>13</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2.5</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2.5</td>
<td>6.5</td>
<td>2.5</td>
</tr>
<tr>
<td><strong>6.4 ± 2.6</strong></td>
<td><strong>7.2 ± 3.5</strong></td>
<td><strong>7.4 ± 2.3</strong></td>
<td><strong>8.2 ± 3.8</strong></td>
</tr>
</tbody>
</table>
Figure 4.7: Boxplot displaying the subject variation of each group, in terms of difference in ROM values obtained before the experimental sessions. The values displayed in degrees correspond to the difference in ROM between the 5th and 1st sessions.

In the project described throughout this thesis, the goal was not to differentiate mental practice from physical practice, nor to compare individuals that already could (somewhat) perform the novel task from those who could not. Our aim was to see if by using a top-down mechanism, subjects would be able to learn more efficiently the novel motor skill, as supported by the central representation theory, than when compared to subjects that only relied on peripheral bottom-up mechanisms based on the activation of muscles (Sham group). Additionally, we expected the Active group to present a bigger increase in ROM values than the Visual group, because the former condition had a physical feedback that would not only correct the movement (allowing for participants to develop a better performance strategy, learning by example) but also close the sensorimotor feedback loop, something that the latter condition did not have.

However, when comparing the values between conditions, no significant difference was found. This might be related to the high subject variability observed, that is accounted for by the low amount of subjects within each group. Even though participants from the Active group significantly increased the physical performance of the movement between sessions, the median ROM values for the sample dataset were similar to the median values collected from the Visual group.

In the postmeasurements, the difference between sessions within a group or even amongst groups was smaller, and no significant change in values was found. This might come from the fact that, after each session, subjects were extremely tired (each session - from capfitting until the end of the main neurofeedback stage - took around 1h45), and several subjects complained about muscle stiffness or soreness. Thus, when the final measurement was done, most participants were not able to perform the target movement with the same vigor as done in the prior ROM measurements.
Figure 4.8: Boxplot displaying the subject variation of each group, in terms of difference in ROM values obtained after the experimental sessions. The values displayed in degrees correspond to the difference in ROM between the 5th and 1st sessions.

4.4 EMG strength

Grand average muscle activity was recorded in this experiment, to observe whether peripheral activity was present and if there was an increase in muscle activation throughout the sessions. In the protocol developed by Mulder and colleagues, only the abductor hallucis muscle of the right foot was measured by resorting to surface electrodes placed on the surface of the belly of the target muscle [68]. However, because this is a novel motor task, we hypothesized that subjects would also activate other muscles, related to big toe flexion, whilst trying to abduct the big toe due to the proximity of the muscles and tendons related to both big toe movements. Thus, the activation of the flexor hallucis brevis and flexor hallucis longus muscles was also recorded with surface electrodes on the right foot and leg, respectively.

The average activity recorded from the three measured muscles is presented, per condition, in Figures 4.9 to 4.11. In these plots, the EMG amplitude was normalized relative to the baseline period of the trials (0-6s), thus the amplitude of the signal is shown in multiples of baseline amplitude (i.e., a value of 1 indicates that the EMG amplitude of that specific timepoint is equal to the median baseline EMG amplitude measured for that muscle, in all trials and subjects of the group). In Figure 4.9, we can observe the normalized EMG activity from subjects belonging to the Active group, in the first and last sessions. All muscles show a much bigger amplitude, after movement onset, in the 5th session when compared to the 1st, with the flexor muscles having a higher activation than the abductor muscle. In order to compare the amplitude values between sessions of the same group, the median amplitude value was computed - within the 7.5 to 19.5s time frame, that corresponds to the movement part of the trial - per subject and session, and a Wilcoxon signed-rank test was applied to this data. Significant differences (p-value = 0.01563 < 0.05) were found between sessions for the values obtained for all three muscles in the
Figure 4.9: Average muscle activity recorded from all subjects belonging to the Active group after the 1st (left) and 5th (right) sessions. The normalized EMG amplitude was computed as multiples of baseline amplitude from the flexor hallucis longus, abductor hallucis, and flexor hallucis brevis muscles. The baseline period was defined as the 0-6s interval, and the red dotted line indicates the beginning of the movement period of the trial, which extends until the end of the time axis.

Active group. The data recorded from subjects of the Visual group, in both sessions, is shown in Figure 4.10. A visual inspection of the plots leads us to conclude that there is no visible difference in average EMG amplitude between the first and last sessions of subjects belonging to this group (although the data measured in the 1st session presents more spikes, probably due to noise that was not eliminated after the processing pipeline). In fact, after applying a Wilcoxon signed-rank test to the data, no significant difference was found for values obtained from participants subjected to the Visual condition, in any of the muscles: the flexor hallucis longus data had a p-value of 0.4688, the abductor hallucis a p-value of 1, and the flexor hallucis brevis a p-value of 0.2969, all higher than the significant level (0.05). Finally, normalized EMG data for subjects of the Sham group is represented in Figure 4.11. The foot muscles flexor hallucis brevis and abductor hallucis show an increase in amplitude, after movement onset, in the last session when compared with the first session, which is not visible for the data measured from the flexor hallucis longus muscle (that maintains a similar amplitude between sessions). However, the median muscle activity data presented p-values of 0.2969, 0.4688 and 0.375, respectively. Thus, no significant difference was found between the 1st and 5th sessions of any of the measured muscles within subjects submitted to the Sham condition.

As was done in previously described performance measurements, the difference in normalized EMG amplitude values, between the 5th and 1st sessions, was computed based on the median muscle value obtained during the movement part of the trials. A Kruskal-Wallis test was applied to this data, in order to compare the performances between the conditions for each muscle. A significant difference was found between groups, for muscle data measured from the flexor hallucis longus ($\chi^2 = 11.963$, p-value = 0.002525 < 0.05). Figure 4.12 displays the distribution of the difference values of this muscle, for each group, in the form of boxplots. We can observe that the Active condition shows the highest amplitude difference values, with a median increase of around 24 times baseline amplitude, between the first and last session. Subjects that belong to the other two conditions, show lower median difference values and, also, less subject variability within the group. The Visual condition shows a median increase between sessions of about 4 times baseline amplitude, however the Sham group shows a negative median value (around -1) indicating that, on average, there was a decrease in normalized EMG amplitude, when comparing the 1st
Figure 4.10: Average muscle activity recorded from all subjects belonging to the Visual group after the 1st (left) and 5th (right) sessions. The normalized EMG amplitude was computed as multiples of baseline amplitude from the flexor hallucis longus, abductor hallucis, and flexor hallucis brevis muscles. The baseline period was defined as the 0-6s interval, and the red dotted line indicates the beginning of the movement period of the trial, which extends until the end of the time axis.

and 5th sessions, for the data recorded from the flexor hallucis longus muscle. A post-hoc pairwise Mann-Whitney U-test was applied to the data to determine which pairs of independent samples presented (or not) a same distribution. Significant differences were found when comparing EMG amplitude difference data between the Active and Visual conditions (p-value = 0.007 < 0.05), as well as the Active and Sham conditions (p-value = 0.022 < 0.05). No significant difference was found when comparing data from the Visual and Sham conditions (p-value = 0.053 > 0.05). Figure 4.13 represents the distribution of difference values for the data measured from the abductor hallucis muscle, also in the form of boxplots. This data has a big subject variability within all conditions, however the Active group still presents the highest median amplitude difference between sessions (around 12 times baseline amplitude), followed by the sham group (median of around 1) and then by the Visual group (which shows a negative median value of -2.5 , indicating that - on average - there was a decrease in normalized EMG amplitude from the first to the last session, for subjects belonging to this group). For the data obtained from this muscle, no significant difference was found between groups, after applying a Kruskal-Wallis test ($\chi^2 = 3.6141$ , p-value = 0.1641 > 0.05). Finally, the boxplots representing the distribution of the difference in EMG amplitude between the 5th and 1st, for data measured from the flexor hallucis brevis, are displayed in Figure 4.14. In this case, all of the conditions present median difference values above 0, indicating that on average there was an increase in muscle activation between the first and last sessions, with the Active group displaying the highest median difference in amplitude (18), followed by the Sham group (5), and the Visual group (2.5). The latter has the lowest subject variability, but the highest skewness in data, towards values that are above the median. As occurred with the abductor hallucis data, no significant difference between conditions was found for this EMG data ($\chi^2 = 3.4954$ , p-value = 0.1742 > 0.05).

The EMG results obtained were in accordance with what was hypothesized prior to the experiment: the participants that were subjected to the active neurofeedback paradigm, and thus received physical feedback related to their brain modulation, showed a significant increase in performance between the first and last experimental sessions. This did not occur for subjects belonging to the Visual or Sham groups, which leads us to conclude that not only is the physical correction of the movement important for the increase in muscle activation, when trying to learn how to perform a new motor movement,
Figure 4.11: Average muscle activity recorded from all subjects belonging to the Sham group after the 1st (left) and 5th (right) sessions. The normalized EMG amplitude was computed as multiples of baseline amplitude from the flexor hallucis longus, abductor hallucis, and flexor hallucis brevis muscles. The baseline period was defined as the 0-6s interval, and the red dotted line indicates the beginning of the movement period of the trial, which extends until the end of the time axis.

but that the active engagement in the task at hand is also a key element for this increase in muscle activity throughout sessions. Previous studies that resort to neurofeedback training with BCIs, involving a mechanical orthosis for correction of the desired movement, also presented improvement in motor function and/or voluntary EMG activity between the first and last training sessions [112, 116].

However, this increase in muscle activity was shown for all three measured muscles of the Active group, and not only for the abductor hallucis that is the target muscle related to big toe abduction. The lack of specificity in the muscle activation leads us to believe that, although the physical and neurofeedback were important for learning how to perform the novel task, it was not sufficient for the development of fine motor control. In fact, when comparing the data between groups, the only muscle that showed a significant difference between the Active condition data and the other two conditions, was the EMG measured from the flexor hallucis longus, related to toe flexion. This means that only in the Active group was the activation of this leg muscle significantly higher when compared with the other groups. This problem refers to a relatively common issue in motor learning paradigms: although subjects are instructed to perform a movement in a certain way, there is no control on behalf of the experimenter on the strategies elicited by participants for the performance of the movement. For this experiment, it seems that the subjects from the Active group contracted their leg muscles more than subjects from the other groups, which might be related to the specific type of feedback given in this condition or, simply, a consequence of high subject variability due to the low amount of subjects within each group (that might have dissipated the inter-group difference in values for the other muscles).
Figure 4.12: Boxplot displaying the subject variation of each group, in terms of difference in normalized EMG amplitude values obtained from the muscle *flexor hallucis longus*. The values displayed correspond to the median difference in EMG values, taken from the movement part of the trials (7.5-19.5s), between the 5th and 1st sessions.

Figure 4.13: Boxplot displaying the subject variation of each group, in terms of difference in normalized EMG amplitude values obtained from the muscle *abductor hallucis*. The values displayed correspond to the median difference in EMG values, taken from the movement part of the trials (7.5-19.5s), between the 5th and 1st sessions.
Figure 4.14: Boxplot displaying the subject variation of each group, in terms of difference in normalized EMG amplitude values obtained from the muscle *flexor hallucis brevis*. The values displayed correspond to the median difference in EMG values, taken from the movement part of the trials (7.5-19.5s), between the 5th and 1st sessions.
5. Conclusion

Stroke is one of the leading causes of long-term motor disability amongst adults. Although many approaches have been used to try to improve motor function recovery, few standard interventions show results that completely restore muscle control. Consequently, patients that are still left with some degree of functional impairment cannot reach full independence in their daily lives. BCI systems have the potential to become part of a new generation of neurorehabilitation techniques for patients with severe paresis after stroke, due to their inherent properties that might facilitate neural network plasticity and restore function through motor relearning.

The aim of this dissertation was to investigate how BCI training can influence neuroplasticity during the acquisition of a new motor skill. Most of the previous research regarding the use of BCI in rehabilitation has focused on upper arm rehabilitation. However, for stroke patients the use of the lower extremities, and more specifically gait, is an important factor in becoming independent of the care of others. Thus, we intended to analyze the effects of neuromodulation of SMR, during the process of lower limb motor learning. For this, we tested the effectiveness of a BCI training protocol, that relied on the use of a robotic device to give physical feedback to healthy participants, in learning how to perform toe abduction.

Behavioral and electrophysiological data was collected from a set of 21 participants, divided into three different conditions - active robotic feedback, sham robotic feedback and visual feedback. Four different types of outcome measurements were analyzed: classifier performance rate, sensorimotor rhythm strength (i.e., ERD effect), EMG strength and ROM values. We hypothesized, prior to the experiment, that the Active and Visual groups would show an increase in classifier performance and a stronger ERD effect, throughout sessions, due to the modulation of cortical excitability caused by the contingent BCI feedback. Additionally, we speculated that the Active and Sham group might display an increase in ROM values and EMG strength, as a result of the physical training of the abductor hallucis muscle. Thus, the overall supposition was that participants that belonged to the Active group would show an increase in performance for all types of measurement, throughout the five experimental sessions, proving that the active control of a robotic orthosis promotes the learning of a novel motor skill.

The results obtained were not as straightforward as we anticipated. Although the average classification accuracies were above chance level (50%), for both the discrimination of the 'Rest' vs. 'Toe Abduction' and for the 'Toe Flexion' vs. 'Toe Abduction' tasks, no significant differences were found between the three conditions nor within the same condition (when comparing the first and last training sessions). This indicates that the logistic regression classifier used in the experiment was able to distinguish the abovementioned tasks, which is further verified by the visual inspection of the amplitude-spectra and AUC plots that display spectral differences between classes in the $\alpha$ and $\beta$ frequency bands over sensorimotor channels. However, the CR values obtained were not consistent within each group, probably due to the high subject variability that came from the lack of subjects per-group. Additionally, despite the processing steps taken when training the classifier, EMG,
EOG and movement artifacts are present in peripheral channels. This artifact corruption of the training data might have lead the classifier to pick up information from broadly distributed channels and, thus, affect the classification accuracy.

Similarly to the CR values, no significant differences were found intra or intergroup-wise when observing the ERD/ERS values computed from sensorimotor oscillations in the $\beta$ frequency range. However, the TFR and topoplots show that for all conditions there is a decrease in power occurring predominantly over centro-medial areas, characterized by a suppression of the $\beta$ and (with less highlight) $\alpha$ frequency bands. This effect is more prominent for subjects that belong to the Active group, suggesting that the coupling of proprioceptive and visual online feedback has a more effective closing of the sensorimotor loop, and leads to a stronger modulation of the signal. However, the fact that there were no significant differences between groups indicates that the neuromodulation of subjects from the Active condition was not strong enough to distinguish them from participants of the other two conditions. Once again, this might be explained by the low amount of participants within each group, which entails a bigger impact of subject variability on the results. The SMR modulation displayed by the Sham group was not expected, seeing as they received mock feedback. Nonetheless, all subjects went through an initial calibration session where they had to attempt to perform each task correctly and the Sham participants were not aware that they were receiving mock feedback. Thus, even without getting explicit feedback, they were still trying to perform the task correctly. Additionally, they were subjected to passive movements of the toe by the robotic apparatus, and this afferent excitation could also have contributed to the frequency changes, that were similar to the ones observed in the active neurofeedback conditions. The non-significant differences between the 1st and 5th sessions can be a possible consequence of the lack of motivation observed in the participant pool. In order to promote motor learning of the novel skill, and to obtain enough data for analysis, we had to design the task in a way that participants repeated the same movement several times throughout the session. When inquired, participants stated that the task was tedious and often they felt distracted and drowsy. Motivation is a key influencer of neural dynamics, and could account for the results obtained. An interesting finding was the difference in spatial distribution of $\beta$ suppression between groups. The Active group showed a lateralization of the ERD effect, towards the left hemisphere, that became more pronounced in the last session, in comparison with the first session. This distribution seems to also occur for the Sham group (with much less intensity and consistency), and not at all for the Visual group. This topographical difference might be a consequence of the design of the physical feedback device.

This last report was further verified by the EMG results. The data measured from the flexor hallucis longus muscle of subjects of the Active condition was significantly different from the same data acquired from subjects of the other two conditions. This justifies the observed lateralization of the signal - only the Active condition participants elicited performance strategies that lead to a stronger contraction of the right leg muscles, and subsequently evoked a $\beta$ desynchronization leaning towards the left hemisphere. Although this was the only EMG signal that showed a significant difference between groups, when comparing the three measured muscles between the 1st and 5th sessions, all muscles showed a significant increase in signal amplitude within the Active group. No significant increase throughout sessions was found for the Visual or Sham conditions. Thus, the physical correction of the target movement, coupled to the active engagement in the task, was relevant for the increase in muscle activation between training sessions.

The ROM values measured prior to the experimental sessions also showed a significant increase, between the first and last sessions, for participants of the Active condition. This behavioral data was our primary outcome measurement of motor learning, hence the fact that only the Active group significantly
increased the angle of toe abduction between sessions allows us to state that, in fact, this was the only condition that learned how to perform the novel task. However, when comparing this data intergroup-wise, no significant difference was found which might be, once again, a consequence of the strong effect of subject variability (that might have dissipated the intergroup difference in ROM).

In summary, although the acquired brain activity data was not statistically different between groups or sessions, the muscle and behavioral data showed an improvement in performance for the Active group, between sessions, serving as proof of concept for the usability of BCI training for the learning of a novel skill. However, we have to be cautious on the drawn conclusions because, for most of the obtained measurements, we did not verify a significant difference between groups. The visual inspection of the computed figures suggest that there is a tendency for a stronger modulation of SMR when there is an active control of the device, but better performances could be achieved if a bigger subject sample is used. Improved signal processing methods - to remove the EMG and movement artifacts from the recorded EEG signals, without enhancing the computational load put upon the system - should also be implemented in future work, to avoid signal corruption by such artifacts (that can be a confound in the differentiation of classes). In addition, a more appealing paradigm is desirable to promote engagement in the task at hand - a fundamental trait, linked to the success of this type of feedback systems. Further studies are needed to confirm the effectiveness of our BCI training protocol and clarify the neural mechanisms related to lower limb motor learning.
Bibliography


Appendix

Figure A.1: Illustration of the intrinsic muscles located within the right leg and foot, taken from Seeley et al. [117]. Muscle names are in bold font, and the referred ones are responsible for the flexion, extention, abduction, and adduction of the toes and foot. (a) Superficial muscles of the right foot, where we can find the *abductor hallucis* inserted in the base of the proximal phalanx of the great toe. This muscle is responsible for the abduction of the great toe. (b) Deep muscles of the right foot, including the *flexor hallucis brevis* inserted in the two tendons near the proximal phalanx of the great toe. This muscle allows for the flexion of the great toe. (c) Posterior view of the deep muscles of the right calf, where we can observe the *flexor hallucis longus* inserted in the distal phalanx of the great toe. This deep muscle of the posterior compartment plantar flexes and inverts the foot, as well as flexes the toes.
Figure A.2: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s2 of the Active group. The magnitude of the spectral differences between 'Toe Abduction' and 'Rest' classes are indicated through the colorbar.
Figure A.3: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s9 of the Active group. The magnitude of the spectral differences between “Toe Abduction” and “Rest” classes are indicated through the colorbar.
Figure A.4: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s5 of the Visual group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Rest’ classes are indicated through the colorbar.
Figure A.5: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s7 of the Visual group. The magnitude of the spectral differences between 'Toe Abduction' and 'Rest' classes are indicated through the colorbar.
Figure A.6: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s17 of the Sham group. The magnitude of the spectral differences between 'Toe Abduction' and 'Rest' classes are indicated through the colorbar.
Figure A.7: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s26 of the Sham group. The magnitude of the spectral differences between 'Toe Abduction' and 'Rest' classes are indicated through the colorbar.
Figure A.8: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s2 of the Active group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Toe Flexion’ classes are indicated through the colorbar.
Figure A.9: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s9 of the Active group. The magnitude of the spectral differences between 'Toe Abduction' and 'Toe Flexion' classes are indicated through the colorbar.
Figure A.10: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s5 of the Visual group. The magnitude of the spectral differences between 'Toe Abduction' and 'Toe Flexion' classes are indicated through the colorbar.
Figure A.11: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s7 of the Visual group. The magnitude of the spectral differences between 'Toe Abduction' and 'Toe Flexion' classes are indicated through the colorbar.
Figure A.12: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s17 of the Sham group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Toe Flexion’ classes are indicated through the colorbar.
Figure A.13: Per-class amplitude-spectra plot (on the left) and AUC plot (on the right) obtained after preprocessing the calibration data from all sessions of subject s26 of the Sham group. The magnitude of the spectral differences between ‘Toe Abduction’ and ‘Toe Flexion’ classes are indicated through the colorbar.
Figure A.14: TFR plots, with the averaged data for each session of the Active group (from first to last). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between rest and abduction, and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis.
Figure A.15: TFR plots, with the averaged data for each session of the Visual group (from first to last). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between rest and abduction, and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis.
Figure A.16: TFR plots, with the averaged data for each session of the Sham group (from first to last). On the right, all channels are represented, on the left the focus is on the central Cz electrode. The TFR plots are relative to baseline, thus 1 indicates that there is no change between rest and abduction, and the red and blue colors imply that there was a relative increase or decrease in activity, respectively. The trial starts with the non-movement baseline period, which lasts 6s. The red dotted line indicates the beginning of the movement period, which extends until the end of the time axis.