Investigating the possibilities of using singing imagery to enhance EEG-based active BCIs

by

© Hadi Mohammadpour

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Engineering.

Department of Electrical & Computer Engineering

Memorial University of Newfoundland

May 2023

St. John's, Newfoundland and Labrador, Canada

I would like to dedicate this thesis to my family,

to the memory of my sister, who always encouraged me to learn from my failures to do better in the future,

to the memory of my mother, who inspired me to push myself and work hard to change the world for the better,

to my father, who, for me, is the epitome of patience, kindness, and love,

to my brother, who has been there for me every step of the way as a great mentor and a role model,

and to the memory of my grandmother, whom I dearly miss.

Abstract

Active brain-computer interfaces are a novel communication/interaction pathway that relies on the performance of imagery tasks. The patterns of brain activity associated with these tasks are detected and decoded as commands for operating an external device. This study focuses on augmenting the practicality of such systems by investigating the mental task of singing imagery. Singing imagery is the simple act of imagining singing a song in your head. Despite its straightforward nature, the potential of singing imagery as an alternative task for active BCIs or for increasing their number of commands has yet to be thoroughly investigated.

The research described in this thesis comprises two phases. In the first study, singing imagery is combined with the commonly used imagery tasks in BCI research (i.e., 4- and 5-class combinations consisting of the imagined movement of the left hand, right hand, feet, and tongue, as well as a "rest" state). Filter bank common spatial patterns algorithm and the random forest classifier are utilized to incorporate a singing imagery task in the 2-, 3-, 4-, and 5-class combinations. These analyses resulted in comparable classification accuracies to conventional motor imagery tasks. Hence, based on the survey results, singing imagery could be considered as a potentially more intuitive alternative mental task. Furthermore, singing imagery may also be a practical approach for increasing the number of commands to six, where accuracies as high as 60.7% were achieved.

The second study investigated the potential of using "dual imagery" tasks (i.e., the simultaneous performance of two single tasks, in this case, singing imagery and one of the conventional motor imagery tasks) as additional BCI control tasks. Here, the 3- and 4-class analyses of the dual tasks and their constituent single tasks (alongside a "rest" state for the 4-class) were carried out to verify the possibility of differentiating them. Using an extended version of filter bank common spatial patterns and regularized linear discriminant analysis classifiers, average accuracies as high as 64.1% and 63% were achieved for the 3, and 4-class scenarios, respectively. Next, the dual imagery tasks were combined with conventional single motor imagery tasks to investigate increasing the number of commands to seven or eight. As a result, for the 7- and 8-class scenarios, accuracies as high as 55.4%, and 50.5%, which are well above the corresponding chance levels of 14.3% and 12.5%, were obtained.

Increasing the number of commands a BCI can recognize is important as it can significantly impact the user's experience with the device. Specifically, a BCI with a more intuitive list of commands can help the user avoid a high mental workload. Moreover, a higher number of commands can be helpful by allowing users to communicate with a higher information transfer rate. Based on the results of this thesis research, singing imagery appears to be a potentially viable solution for improving active BCIs.

Acknowledgements

I would like to sincerely thank my advisor Dr. Sarah D. Power for her support and valuable guidance during the course of my research. I will forever be grateful for having the opportunity to work with her on this project.

I would also like to thank all my friends and colleagues at Memorial University of Newfoundland including but not limited to Dr. Mahsa Bagheri, Mr. Jacob White, Ms. Faghihe Massaeli, Mr. Roy Brushett, and Ms. Mona Hejazi for their help and support.

I would also like to acknowledge the financial support of the Natural Science and Engineering Research Council of Canada (NSERC) Discovery Grants Program.

Finally, and most importantly, I would like to thank my family for their ever-present support of my ambitions.

Table of contents

List of tables

List of figures

List of abbreviations

BCI	Brain-Computer Interface
CNS	Central Nervous System
MI	Motor Imagery
R	Right Hand Motor Imagery
L	Left Hand Motor Imagery
F	Feet Motor Imagery
T	Tongue Motor Imagery
SI	Singing Imagery
SI_{Kin}	Kinesthetic Singing Imagery
SI_{noKin}	Non-Kinesthetic Singing Imagery
DI	Dual Imagery
R_{SI}	Dual Right Hand Imagery
L_{SI}	Dual Left Hand Imagery
F_{SI}	Dual Feet Imagery
ALS	Amyotrophic Lateral Sclerosis
EEG	Electroencephalography
MEG	Magnetoencephalography
fMRI	Functional Magnetic Resonance Imaging
fNIRS	Functional Near-Infrared Spectroscopy
FIR	Finite Impulse Response
CAR	Common Average Re-referencing
CSP	Common Spatial Pattern
FBCSP	Filter Bank Common Spatial Pattern
mRMR	Minimum Redundancy Maximum Relevance
KNN	K-Nearest Neighbors
SVM	Support Vector Machine
LDA	Linear Discriminant Analysis
RLDA	Regularized Linear Discriminant Analysis
ICEHR	Interdisciplinary Committee for Ethics in Human Research
KVIQ-10	Kinesthetic and Visual Imagery Questionnaire

Chapter 1

Introduction

Portions of this chapter have been submitted as part of the following manuscripts:

- "Investigating Singing Imagery as a Potential Control Task for Motor Imagery BCI", which is currently under review at IEEE Transactions on Biomedical Engineering (Manuscript ID: TBME-00207-2023).
- "Investigating Dual Imagery Tasks for BCI Control", which is currently under review at Journal of Neural Engineering (Manuscript ID: JNE-106478).
- "Investigating the addition of singing imagery as a control task in motor imagery BCI", which is a conference proceeding presented at IEEE SMC 2022 conference in Prague, Czech Republic. [43]

1.1 Brain-Computer Interfaces

Brain-computer interfaces (BCI) are a recent technology that allows for direct communication between the brain and external devices. A BCI is defined as *"A system that measures central nervous system (CNS) activity and converts it into artificial output that replaces, restores, enhances, supplements, or improves natural CNS output and thereby changes the ongoing interactions between the CNS and its external or internal environment."* [67] In particular, "active BCIs", which are the focus of this work, provide users with a movement-free pathway to communicate and interact with their surroundings. The user controls the BCI by intentionally generating a predefined set of brain activity patterns which are detected and translated into commands for an external device (e.g., a computer). This is in contrast to "passive BCIs", which are not consciously controlled by the user but rather derive explicit information about the user's state (e.g., fatigue, attention) and use it to enhance some human-computer interaction.

1.2 Mental tasks in active BCIs

For an accurate and efficient active BCI design, the generated EEG signal patterns should be clear, distinct, and reproducible. Often, various mental tasks are used to help the user generate these patterns, with each task being associated with a different command.

Currently, the most common type of mental task used for active BCIs is the kinesthetic motor imagery (MI) of body parts. Kinesthetic motor imagery refers to mentally simulating a movement by recalling the sensations associated with the actual execution of the movement [48] (this is different from visual motor imagery, which refers to visualizing a movement without recalling the associated sensations). Interestingly, this mental simulation of movement generates similar patterns of brain activity, with the same cortical areas involved in both [45, 24, 68]. The most commonly used tasks in EEG-based MI-BCI research are the imagined movement of large body parts like the hands, feet, and tongue. However, some users may find these tasks difficult to perform, and or struggle with generating clear patterns of neural activity that can be detected by the BCI algorithm. This may be particularly true for users with severe motor disabilities who may not have actually executed the associated movements in a long time (or, possibly, ever).

Alternatively, some users may find non-motor tasks more suitable. Although such tasks are not as well-studied as MI tasks, some non-motor tasks such as mental calculation/counting [46, 3, 10, 57, 12, 18, 55, 9, 16, 15, 47], word generation/association [46, 3, 10, 57, 12, 18, 55, 9, 16, 47], spatial navigation [3, 18, 55, 19, 16, 15, 7, 17], mental object rotation [3, 10, 12, 18, 55, 9], and facial imagery [18, 55] have been explored in the BCI literature. Unfortunately, however, some users may find them difficult to perform. Even if high classification accuracies could be achieved with these tasks in a research setting with healthy participants, they may not result in practical, user-friendly BCI solutions for frequent and long-term use by the target population of users. Thus, **a thorough investigation of alternative, intuitive non-motor tasks for designing practical active BCIs is needed.** This investigation should evaluate the alternative task both in terms of performance and also perceived difficulty.

1.3 Increasing the number of commands

As mentioned, in active BCI systems mental tasks are often used to help the user generate distinct patterns of brain activity that can then be associated with different output commands. Therefore, increasing the number of mental tasks that the BCI can detect would result in increasing the number of possible commands, thereby augmenting the functionality and practicality of the system. Indeed, increasing the number of supported commands has been noted as a key issue in the field of active BCI [69]. In current EEG-based BCI research, up to four motor tasks/commands are typically investigated, specifically motor imagery of the left hand, right hand, both feet, and tongue. In terms of non-motor tasks, a majority of studies have focused only on binary classification between pairs of tasks [46, 3, 10, 57, 12, 18, 55, 7], while a few have explored classification of 4-classes [9, 19, 15, 47].

A relatively small number of MI-BCI studies have investigated the classification of more than four tasks. Christensen et al. [6] explored the possibility of controlling a drone using a 5-class MI-BCI where a rest state was added to the conventional tasks of R, L, F, and T motor imagery. This 5-class paradigm resulted in an average classification accuracy of 41.8% in an online study with 10 participants. One study investigated the possibility of a 5-class BCI based on the four common MI tasks (i.e., L, R, F, and T) and a mental calculation task [47]. This combination of five tasks yielded a mean accuracy of 53% across three participants. Moving away slightly from the conventional tasks, Faiz et al. [11] proposed using different imagined movements of a single hand (i.e., open, close, pronate, and supinate) along with a rest state to develop a 5-class system. They reported a high classification accuracy of 97.5% with a single participant; however, it is not clear whether such levels of performance would be maintained for a larger sample size.

Very few studies have looked beyond a 5-class MI-BCI system, and those that have did not use additional MI tasks but rather proposed paradigms in which a lower number of MI tasks were used in sequence to achieve multiple commands [69, 30, 66]. Focusing on the conventional tasks of L, R, and REST, Jiang, et al. [30] used pairwise combinations of these three tasks, performed in sequence, to produce a total of six different commands (i.e., L-R, L-REST, R-L, R-REST, L-L, and R-R). Using this paradigm, an average accuracy of 89.4% was obtained across four participants. Unfortunately, though, even if these commands can be distinguished with high accuracy, the difficulty of performing sequential tasks, along with the low information transfer rate inherent in such multi-step command paradigms, may potentially cause user frustration and abandonment of the device.

Increasing the number of commands should not undermine the intuitiveness of BCI control from the perspective of the user. As mentioned earlier, the target users of active BCI systems are often individuals with severe physical disabilities, and these individuals may find it difficult to go through long and demanding training sessions to learn how to control the BCI. Therefore, the defined tasks ought to be relatively easy to perform, and also reasonably different from each other to avoid confusion. For example, the intrinsic similarities among the four single-hand tasks proposed in [11] (i.e., open, close, pronate, and supinate) may lead to difficulties for users with severe motor disabilities who may not have performed such precise movements in a very long time. Hence, **exploring potentially effective and user-friendly approaches for increasing the number of commands in active BCIs is justified.**

1.4 Singing Imagery (SI) as a potential control task

Singing imagery (SI) is an intuitive task that, despite its potential advantages, is yet to be thoroughly examined for use in active BCIs. Performing singing imagery is as simple as imagining singing a song in your head, an experience that is very common to most people. Because of the familiar and intuitive nature of the task, SI has the potential to simplify BCI control from the perspective of the user. Along with issuing discrete commands, SI may be particularly advantageous for applications in which a desired action must be maintained continuously for an unspecified period of time; for example, when scrolling through content on a computer screen or moving a video game avatar forward. Performing most of the other mental tasks mentioned above (both motor and non-motor) for more than a few seconds at a time would be quite difficult. On the other hand, it would be relatively easy for a user to simply start and then continue performing singing imagery to scroll up or down on a computer screen until they reach their desired content or to move a video game avatar forward until they wish it to stop.

A few EEG-BCI studies have included singing imagery as one of a longer list of investigated mental tasks, but it was only considered in binary classification scenarios, and the accuracies reported were generally low in comparison to other tasks [46, 3]. However, these low accuracies may have been due to how the singing task was defined in these studies, where participants were instructed to simply "sing a song of their choice in their heads" [46], or to "imagine singing a song that they chose beforehand, if possible with lyrics, while focusing on the emotional response it elicits" [3].

It is reasonable to think that the brain activation associated with singing imagery will be most clear and distinct if it is treated as 1) a special case of mental speech (in which cortical activation is said to be due to articulation preparation, including motor planning, as well as auditory cortex activation caused by efference copies [49]) and/or 2) a motor imagery task itself (i.e., imagined movement of the tongue, lips, and jaw). Therefore, it is likely that having the user focus on the imagery of articulating the lyrics of the song while performing singing imagery will be critical to generating brain activity patterns that can be reliably detected. However, none of the previous studies that considered singing imagery as a BCI control task explicitly instructed the participants to do this.

It is worth noting that although in [3], singing imagery had relatively poor classification performance, it was also rated lowest of all seven mental tasks investigated in terms of mental demand, effort, and frustration. Thus, given its potential advantages, **a thorough investigation of the potential of singing imagery as a BCI control task - in both binary and multi-class scenarios and with a focus on the speech/motor aspects of the task - is warranted.** This was the aim of Study 1 of this thesis research.

1.5 Dual imagery (DI) as potential control tasks

A mental task involving the simultaneous performance of two usually distinct mental tasks can be termed a "dual imagery (DI)" task. For instance, motor imagery of the left hand and the right hand, which are differentiable from one another by EEG, could be performed simultaneously to produce a third task. Although there is some disagreement as to the effects of multitasking on the brain [1, 28], most agree that combining several low-level, simplified tasks has an additive effect [25]. Therefore, it is reasonable to suggest that combining two individual tasks would produce a third DI task with a distinct pattern of brain activity, such that all three tasks would be differentiable from one another. If so, using such DI tasks along with single tasks could be a novel approach to increasing the number of classes a BCI can detect and, thus, the number of BCI commands. Detecting such dual tasks could be particularly useful in applications like wheelchair/cursor control where the single tasks could be translated into movement in the four cardinal directions (i.e., north, east, south, and west), and their dual task could translate into movement in the corresponding ordinal direction (i.e., northeast, southeast, southwest, and northwest) [33].

Very few previous BCI studies have investigated DI tasks. One study attempted to classify DI tasks in a 4-class scenario consisting of "no motor imagery", "imagery of left hand", "imagery of right hand", and "imagery of both hands" [21]. Based on a limited sample size of participants (n=3), promising results were reported for this paradigm, with a maximum accuracy of 75.8% across the participants. However, the potential of DI for increasing the number of commands beyond four was not investigated.

The possibility of using DI tasks to increase the number of commands from four to five was explored in Reshmi et al. [54], where the simultaneous performance of R and L replaced the T imagery task. Moreover, Kim et al. [33] also combined two motor imagery tasks to create a separate command; specifically, the simultaneous performance of R + F and L + F were used with the single tasks of R, L, and F to build a five-class system. Nevertheless, neither of these two studies reported performance evaluation metrics that can be used to gauge the potential functionality of the proposed protocols.

Another study on DI tasks investigated the eight tasks that could be achieved by combining motor imagery of the left and right arms and both feet [62]. However, despite recording EEG data from 18 participants, they reported classification results for only three of them who had distinguishable patterns (based on the BCI algorithm used in their study). Accuracies as high as 50% were reported for the 8-class classification across the three participants.

The investigated DI tasks, generated by two MI tasks, tend to be tiring and demanding for the brain. Moreover, it can be argued that due to the inherent similarities of different MI tasks and their corresponding patterns, their corresponding dual task may not provide desirable levels of performance. Alternatively, the act of singing a song and tapping your hand or foot along to the rhythm is a very common experience for most people, so this could produce a relatively more intuitive and easier DI task combination. Furthermore, the patterns associated with SI and MI may be more differentiable in comparison to two MI tasks. Hence, **an investigation of DI tasks created by the simultaneous performance of motor imagery of limbs and singing imagery is justified.** Here, based on the findings of Study 1, which investigated the potential of singing imagery as a BCI control task and verified the differentiability of SI and MI tasks, Study 2 investigated DI tasks composed of a conventional motor imagery task performed simultaneously with singing imagery.

1.6 Research objectives

The overall objective of this study was to examine the potential of the novel mental task of singing imagery to enhance the current state-of-the-art active BCI systems. In particular, the main goal was to investigate the potential of using singing imagery as an alternative to the traditional four tasks - motor imagery of the left hand (L), right hand (R), feet (F), and tongue (T) - as well as in combination with these tasks to increase the number of possible BCI commands above four. Thus, the specific research objectives of this thesis were as follows.

1. Considering the following set of mental tasks (SI, L, R, F, T, REST):
   (a) Investigate the accuracy with which all 2-, 3-, 4-, and 5-class combinations can be classified via EEG, and compare the combinations including SI with those including only conventional MI tasks.
   (b) Determine the accuracy with which all six tasks can be classified via EEG.
   (c) Compare SI to the conventional MI tasks in terms of participants' subjective preferences and perception of task difficulty.
2. Considering the following set of mental tasks (L, R, F, SI, L_{SI}, R_{SI}, F_{SI}, REST), where L_{SI}, R_{SI}, and F_{SI} are dual imagery tasks:
   (a) Investigate the accuracy with which all 3-class combinations consisting of a dual imagery task and its two constituent single tasks (e.g., L_{SI}, L, and SI) can be classified via EEG.
   (b) Investigate the accuracy with which all 4-class combinations consisting of a dual imagery task, its two constituent single tasks (e.g., L_{SI}, L, and SI), and REST can be classified via EEG.
   (c) Investigate the accuracy with which all 7-class combinations can be classified via EEG.
   (d) Investigate the accuracy with which all eight tasks can be classified via EEG.
   (e) Compare DI tasks to single tasks in terms of participants' subjective preferences and perception of task difficulty.

1.7 Thesis organization

The rest of this thesis is organized as follows:

Chapter 2 provides a thorough review of the literature relevant to the current thesis, including a discussion of active BCIs and their design. The main components of active BCIs are described, and the specific machine learning and signal processing algorithms and techniques used in this research are discussed in detail.

Chapter 3 presents the first study conducted as a part of this thesis in which the potential of using singing imagery as a control task in active BCIs was explored. The study design, methodology, results, discussions, and conclusion of this study are presented. This chapter addresses research objective #1 in section 1.6

Chapter 4 presents the second study conducted as a part of this thesis in which the potential of using dual imagery (DI) tasks (comprising singing imagery and traditional motor imagery tasks) as control tasks in active BCIs was investigated. The study design, methodology, results, discussions, and conclusion of this study are presented. This chapter addresses research objective #2 in section 1.6

Chapter 5 summarizes the original contributions of this thesis research, briefly discusses the limitations of the work, and recommends future work to be explored.

Chapter 2

Literature review

2.1 Brain-computer interface

Any sort of mental activity leads to an alteration in neurophysiological signals. In brain-computer interfaces (BCI), such changes in brain activity are measured, detected, and translated into commands/information to be sent to an external device. BCIs have led to the development of a novel communication and control pathway that could be used to enhance the human experience with technology or even the quality of a user's life.

Physical movements, speech, and gestures are some of the common ways that people communicate and interact with the world around them. However, for some individuals with severe physical disabilities, such as those arising due to neurological disorders like amyotrophic lateral sclerosis (ALS), these common pathways may not be efficient or effective. BCIs offer these individuals an alternative, movement-free means of communication and environmental control, which is based on intentionally-generated brain activity alone. The goal is to allow users to control an external device, for example, a robotic arm [37], a wheelchair [69], or a computer application [42]. The type of BCI that requires such a direct intention and effort from the user is referred to as "active BCI".

2.2 BCI algorithm

In active BCI systems, the user initiates commands by intentionally generating a particular mental state, often through performing different mental tasks. The BCI uses information contained in the user's neurophysiological signals to decode which mental state the user generated, and thus their intended command, and then outputs the command to a connected external device. To accomplish this, a typical active BCI algorithm comprises the following main stages (see Figure 2.1): 1) neural data acquisition, 2) signal pre-processing, 3) feature extraction, 4) feature selection, 5) classification and 6) control and feedback. The specific methods and techniques used in each stage can vary and will depend on the specific design requirements in a given scenario. These major BCI algorithm components are described in detail in the following sections.

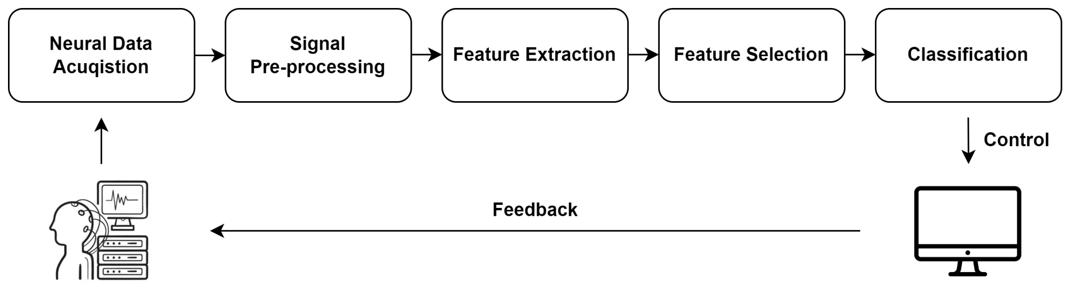

Figure 2.1: Main stages of the typical active BCI algorithm

Note that BCI design generally consists of calibration and online/use phases. Optimizing the feature selection and classification algorithms are typically conducted in the calibration phase, and the selected features and the generated model are then used in the online phase. However, some BCIs utilize adaptive algorithms where the selected features and/or the classification model are periodically updated during the online phase. This can be helpful in addressing changes that occur in the signal characteristics due to factors like fatigue. Also, the feedback only occurs in the online/use phase.

2.2.1 Data acquisition

The first step is to use sensors to record the users' neural signals as they interact with the BCI. This can be done invasively (i.e., sensors are implanted on/in the brain tissue) or non-invasively (sensors are placed on the surface of the scalp). Various recording modalities have been investigated for use in BCIs. These have included non-invasive methods like electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and functional near-infrared spectroscopy (fNIRS) [23], as well as invasive methods like electrocorticography (ECoG) and intracranial electrode arrays. Each of these approaches offers advantages while being limited in some other aspects (see Table 2.1). Therefore, it is important to fully consider the application and choose the recording modality appropriately.

Of the non-invasive technologies, EEG is the one most often used in BCI research. BCI systems must be able to detect rapid changes in neural activity, they must be able to be used in everyday scenarios and settings, and they must be affordable for users, so EEG, with its high temporal resolution [36], portability (the required equipment

Table 2.1: Commonly used recording methods in active BCI systems

Recording Method	Temporal Resolution	Spatial Resolution	Cost	Portable	Invasive
EEG	High	Low	Low	Yes	No
MEG	High	Relatively High	High	No	No
fMRI	Low	High	High	No	No
fNIRS	Low	Relatively High	Low	Yes	No
ECoG	High	High	High	Yes	Yes

is quite small, and there are wireless options available), and relatively low cost is the most promising option to date.

Electroencephalography (EEG)

- **Electrophysiological phenomena**

EEG is a non-invasive neuroimaging technique capable of measuring the firing patterns of billions of neurons [58]. Specifically, the synchronized activity of cerebral neurons, primarily the pyramidal neurons [5], leads to the generation of electrical activities that can be measured via a set of EEG electrodes placed on the scalp.

The electrical fields generated by the neurons which are not aligned in the same direction tend to be very dispersed and, on average, cancel out each other. However, pyramidal cells are oriented perpendicularly to the cortical surface [13] and oscillate at the same frequency. This means that their electrical field is additive and projected toward the scalp and can be captured via the EEG sensors.

Still, it should be noted that due to the poor spatial resolution of EEG signals, even with a highly dense grid of EEG electrodes, it is not possible to detect individual sources of brain activity. This is because the neural activity from different brain regions may be mixed together and measured by different adjacent electrodes, making it impractical to localize the detected activity.

On the other hand, the high temporal resolution of EEG and its inherent spectral properties have proved to be useful in BCI design. Specifically, EEG data is typically broken down into the delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), mu (8–12 Hz), beta (13–25 Hz), and gamma (>25 Hz) components, and each of these bands has been reported to represent different types of neural activation (e.g., relaxation, intellectual activity and focus, and perception and consciousness) [38].

Moreover, in MI-BCI, the imagined movement of body parts leads to a set of patterns known as event-related desynchronization (ERD) and event-related synchronization (ERS) in the EEG data. These patterns are characterized as a decrease or increase of signal power in the brain's somatosensory rhythms, respectively [52], and are typically observed at the sensorimotor area of the brain in the 7 - 30 Hz frequency range [41].

- **EEG recording**

EEG recording systems consist of various components that work together to detect the neural activity, amplify the detected signals, convert the inherently continuous signals to digital format, and transmit and store them on an external device such as a computer. The major components of an EEG system are the electrodes and the amplifier unit, which are connected via cables or wirelessly. The EEG recording setup used in the present work, incorporating the discussed components, is displayed in Figure 2.2.

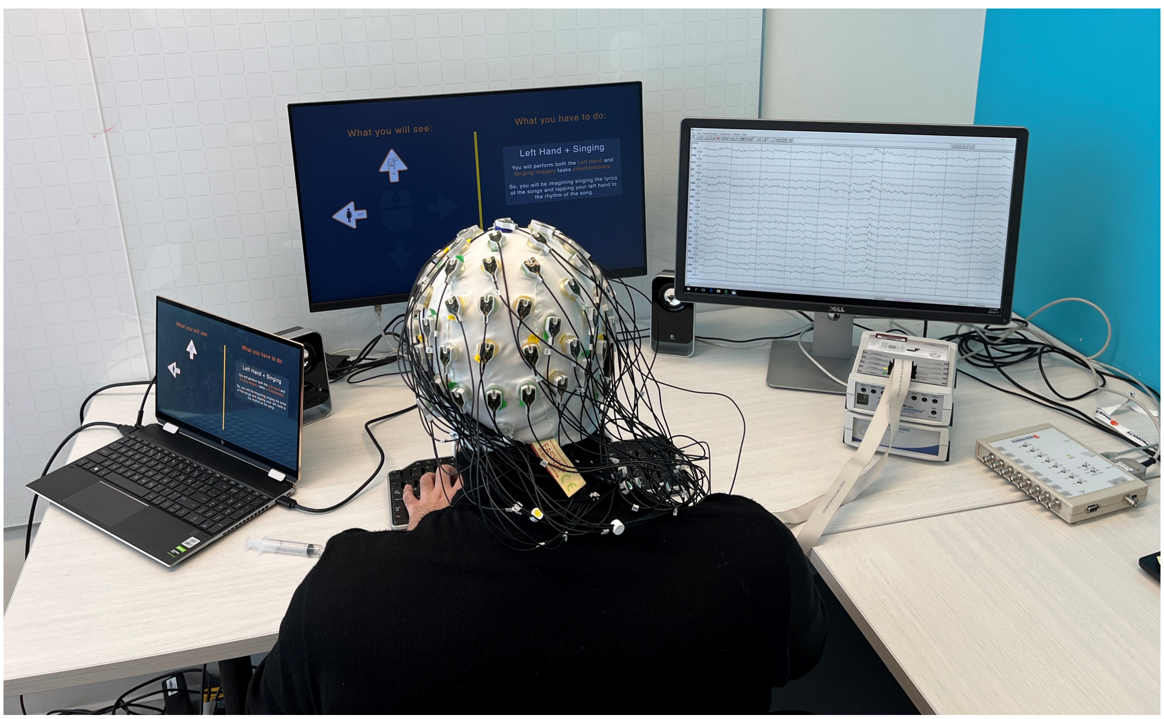

Figure 2.2: EEG recording setup used in the present study

To achieve a high signal-to-noise ratio, a conductive electrolyte gel is typically applied to each electrode to reduce the impedance between the tip of the electrode and the scalp. However, the recent development of practical and comfortable "dry electrodes" could potentially eliminate the need for this in the future [64], which would be a significant advantage for BCI applications. Since the EEG signals detected at the scalp are usually extremely small - in the microvolt range - it is necessary to amplify and enhance these signals for accurate digitization. For the A/D converter, it is important that a sufficiently high sampling rate is chosen so that no information is lost in the conversion from the analog to the digital signal. In MI-BCI studies, sampling rates in the 128 - 512 Hz range are commonly used [63, 51].

EEG recording electrodes can be characterized as active or passive (this is unrelated to the categorization of BCIs as active or passive). Passive electrodes typically have a poorer signal-to-noise ratio than active electrodes. In these electrodes, the signal that is detected at the electrode is transmitted to the amplifier via a wired connection. The signal can pick up noise during transmission, and because the measured EEG signals are extremely small, this can result in a very poor signal-to-noise ratio once the signals are eventually amplified. To mitigate this issue, active electrodes contain circuitry for pre-amplifying the signal immediately as they are detected at the scalp. This minimizes the effect of noise acquired during transmission to the amplifier unit. Therefore, for a practical and portable EEG-based BCI design, active electrodes are a better choice [35].

Another important consideration prior to EEG recording is the placement of electrodes on the scalp. Typically, the standard 10-20 or 10-10 systems adopted by the International Federation of Clinical Neurophysiology are followed [27]. The 10 and 20 in the name refer to the percentage of the total front–back or right–left distance of the skull, which specify the distance between two neighboring electrodes. Figure 2.2 depicts the standard 10-10 electrode placement system for 64 electrodes.

Because EEG is the most widely-used non-invasive recording technology used in active BCIs, and it is the technology used in this thesis research, the following sections will discuss the BCI algorithm with a focus on EEG-based BCI.

2.2.2 Data pre-processing

Once the EEG data is acquired, the first step is to remove noise and artifacts contaminating the signals, as it could affect the performance of the BCI. For that purpose,

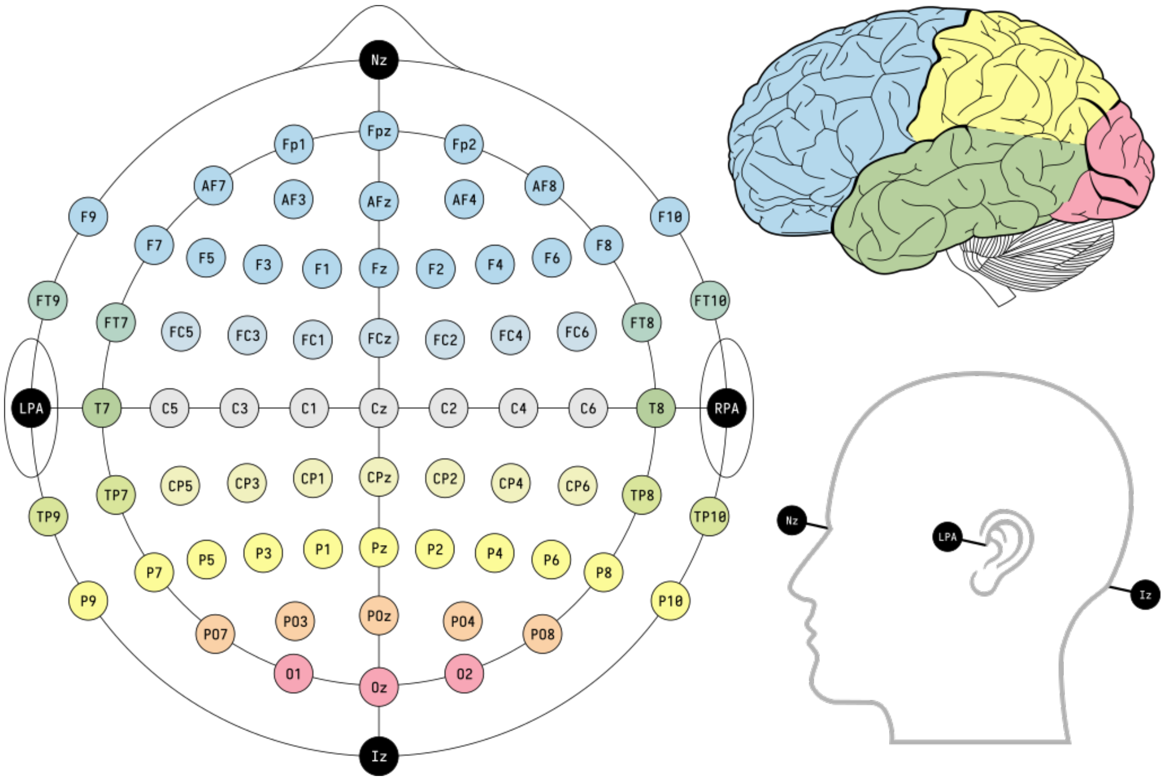

Figure 2.3: The standard 10-10 electrode placement system

the recorded data has to go through a pre-processing pipeline. Some of the typical elements of this pipeline are discussed in the following sections.

FIR filters

Finite impulse response (FIR) filters are a commonly used type of digital filter that manipulate the signals' frequency components in a way that only certain frequencies are allowed to pass, and the rest are attenuated. This behavior of the filter is called the frequency response of the filter. By adjusting the filter's frequency response, one could use it for various purposes, such as noise reduction and extraction of frequency-specific information.

The impulse response of these filters, as the name suggests, is non-zero only for a limited duration after an impulse input. This response is basically a set of coefficients calculated in the filter design process that determine the behavior of the filter. A linear phase response (i.e., the phase delay remains constant across the frequency band), stability (i.e., the output would not grow indefinitely), and ease of design and implementation are some crucial advantages that FIR filters offer.

Due to these qualities, FIR filters are widely used in a variety of applications, like bio-signal processing, image processing, and audio/speech processing. In EEG signal processing, FIR filters are commonly used for removing the DC components of the signal and avoiding the noise-sensitive high-frequency components by bandpass filtering the data in the 1-100 Hz frequency range.

Common average re-referencing

Whilst recording, neural activity is inferred by calculating the relative difference in voltage measured at individual electrodes to the reference point (typically one of the electrodes). Re-referencing is the technique of changing this reference point for the recorded EEG signal. It should be noted that re-referencing can affect the interpretation of EEG data and should be carried out with care.

Common average re-referencing (CAR) is a common approach in EEG signal processing. In this approach, the average of all the electrodes is calculated and then subtracted from the readings of all the electrodes. This can lead to improvements such as the attenuation of common noise sources and the enhancement of spatial resolution.

Epoch extraction

In active BCI design, an epoch is defined as a window of time that the system considers for predicting the mental task that is being performed. In cued systems, epochs are extracted around the period of task execution with a predefined margin before and/or after that period.

It should be noted that parameters such as the window's length, offset before and after the event, and criteria for rejection of epochs should be carefully determined as they can significantly impact the analysis. The chosen parameters should preserve the relevant neural activity and minimize the effect of artifacts in the EEG data.

Baseline correction

Removing the EEG data's DC offset is a commonly used technique in EEG signal pre-processing. The electrode-scalp interface, electrode-cable interface, and amplifier's input offset are some of the sources that can generate a constant bias in the EEG signals. This can negatively impact the signal's interpretation as it could lead to meaningless variations in the data. Removing the baseline and centering the data around zero can also be helpful in enhancing the performance of signal processing techniques that are commonly used in frequency analysis.

It should be noted that applying baseline correction incorrectly can introduce errors in the data. Therefore, an appropriate method should be applied on a valid time window to effectively calculate and remove the DC component without introducing undesired effects.

2.2.3 Feature extraction

At the classification stage of the BCI algorithm, information from the pre-processed EEG signals is used to predict which mental task the user is performing. Typically, the raw signals are not used in this classification, but rather "features" are first derived/extracted from the signals, which contain the salient information needed to differentiate the mental tasks of interest. Features capturing temporal and/or spatial and/or spectral characteristics of the signal are typically useful. Typical time-domain features include signal mean, variance, and standard deviation, while in the frequency domain, the power spectral density of the signal and the band power (i.e., average power over a specific frequency band) are commonly used. To extract spectral features that may vary while performing a specific task, time-frequency methods such as wavelet transform and short-time Fourier transform can be used.

Common spatial pattern (CSP) filters [70, 22] are a type of spatial filter used for extracting relevant information from multi-channel EEG signals. They are the most commonly used feature extraction method in identifying patterns of activity for motor imagery tasks. This approach aims to determine a set of spatial filters that could maximize the separation between two populations of EEG signals (i.e., two mental tasks such as left and right hand motor imagery). In doing so, the variance is maximized for one EEG population and minimized for the other.

Filter bank common spatial pattern (FBCSP) is an extension of the CSP algorithm in which CSP filters are optimized for multiple frequency bands rather than just one [2]. Typically, a bank of FIR filters is used to extract signals in the different frequency ranges of interest. Then, each filter's output is input to the CSP filter optimization algorithm, and the associated features are calculated. FBCSP results in a larger set of features containing more specific spectral-spatial information that is often more useful in differentiating the mental tasks of interest.

2.2.4 Feature selection

The number of extracted features can be quite high in some scenarios, and appropriate feature selection approaches should be deployed to lower this number as it can significantly impact the performance of the system. In particular, a high number of redundant features that do not provide any value may even compromise the system's performance and/or processing time. Moreover, reducing the size of the feature set is also helpful in avoiding the "curse of dimensionality" phenomenon [4], which refers to the exponential increase in the number of training samples required to accurately fit the classifier as the number of input features increases.

There are feature selection algorithms that can automate this process. Feature selection methods can be categorized as wrapper methods, filter methods, and embedded methods. Wrapper methods compare different subsets of features, one at a time, based on their performance with the particular classification algorithm to be used. This solution has a high computation cost and a tendency toward overfitting in cases with lower numbers of training samples. In filter methods, on the other hand, no model is estimated, and the different features are instead evaluated and ranked based on statistical metrics such as variance, correlation, mutual information, etc. Although this approach addresses the issues of wrapper methods, it fails to capture correlations between different features and therefore does not select the optimal feature subsets. Embedded methods aim to find a balance between the two approaches by combining the feature selection approach with the machine learning algorithm. This results in a faster solution that also considers the correlation between different features.

Minimum Redundancy Maximum Relevance (mRMR) [8] algorithm is a filter method that, as the name suggests, aims to identify the features that are most relevant to different classes while avoiding any redundant features. To identify such features, this algorithm computes the relevance of each feature by estimating the mutual information between that specific feature and the class label. However, to avoid redundancy, a feature is only included in the final dataset if the shared information between that feature and the others is minimal. Eventually, a set of features that are highly relevant to the output and are minimally redundant is selected. Due to its potential for improving the performance of classification and regression models, the mRMR algorithm is a popular choice in a variety of applications, such as bioinformatics, text classification, and image recognition.

2.2.5 Classification

In BCIs, machine learning-based classification algorithms are trained to differentiate the various mental intentions/states based on the selected features. Classification is a type of machine learning problem that aims to predict the class membership (i.e., label) of any given unseen input data based on a model trained on data with known labels (i.e., training data). A "decision rule" is specified based on this training data and is used to predict the class of unseen data samples.

Deep learning algorithms are a type of machine learning algorithm in which the features and the classifier are learned from raw data. In other words, these algorithms provide an end-to-end modality that does not require any type of engineered features or representations of data. Several motor imagery studies have investigated the potential of these algorithms [61, 60, 56, 39]. However, despite their capabilities in simplifying and improving BCI design, a very large number of training samples is required for calibrating them. Typically BCIs at the research stage are designed and developed using a small number of samples which are recorded during a limited number of recording sessions. Moreover, the high computational complexity of deep learning algorithms in the training and testing stages requires the usage of high-performing computing tools.

The choice of the classification model is very important and should be carefully considered based on the input data. K-nearest neighbors (KNN), support vector machine (SVM), neural networks, random forest, and linear discriminant analysis (LDA) are common classification algorithms used in BCI research. The latter two algorithms, which were used in this work, will be described in the following sections.

Linear discriminant analysis (LDA)

Linear discriminant analysis (LDA) is a supervised machine learning algorithm. This algorithm aims to identify a linear combination of features in the training data that can be used to categorize the data by maximizing the inter-class variance and minimizing the intra-class variance [20]. This is carried out by projecting the data onto a lower-dimensional space, and the projection is referred to as "linear discriminant". Some of the main features of the LDA algorithm include:

1. It is computationally low-cost and fast, making it a good choice for applications with a large amount of data
2. The simplicity and linear nature of this model allow for higher interpretability. However, this linearity of the LDA algorithm also means that it is not capable of capturing non-linear relationships between features and classes.
3. Using this algorithm comes with an assumption that the data is normally distributed and that the classes share equal covariance matrices, which may be a flawed assumption [29].
4. This algorithm is highly sensitive to outliers.
5. Availability of a large number of features can lead to overfitting of the LDA model.

Regularization is a set of techniques that are used in machine learning to prevent a model's overfitting and underfitting. By incorporating regularization into the LDA algorithm, it is possible to increase the model's bias and reduce the risk of overfitting. For that purpose, a regularization hyperparameter needs to be specified to determine the amount of bias added to the model. This approach is referred to as regularized linear discriminant analysis (RLDA) [14] and aims to improve the generalizability of the model while maintaining a successful fitting of training data.

All these qualities have turned LDA into one of the most popular choices for use in the classification of EEG signals [59, 44, 31, 32].

Random forest

Random Forest, first proposed by Ho [26], is a powerful machine-learning algorithm capable of handling both classification and regression problems. This method is an ensemble method meaning that multiple decision trees come together to generate an output, and the output is not based on just one model. Decision trees are basic classifiers that repeatedly split the data based on certain metrics and criteria with the goal of characterizing different data points as being members of different classes. In the random forest algorithm, a large number of such decision trees vote on the final output, and in doing so, each tree is trained and optimized based on a different subset of features. This randomness improves the performance of the algorithm by reducing the correlation between the trees. For classification problems, the final output is the majority vote of the predictions made by the trees [59, 53].

Some of the main features of the random forest algorithm include:

1. It is a good choice for more complex problems as it is capable of detecting underlying relationships between complex and non-linear features.

2. It can handle large amounts of data and can be run in parallel for faster processing; however, memory usage is high, especially with larger datasets.

3. It can handle missing data (i.e., when some feature values from some samples are missing).

4. It ranks the importance of different features, which can be used for feature selection and simplifying the model.

5. A high number of trees in the forest may lead to overfitting, so the number of trees should be tuned.

6. It is computationally expensive, and the prediction time can be quite long in comparison to algorithms such as LDA.

2.2.6 Feedback / Control

After a prediction is made regarding which mental task the user is performing, the command associated with that task is issued to the external device. The resulting action - whether or not it was the one they intended - serves as feedback to the user regarding how successfully they are controlling the BCI. This may result in the user modifying their performance of the tasks and learning to perform them such that they are more accurately detected. The feedback, of course, also contributes to the user's decision regarding their next command.

2.2.7 BCI performance evaluation

In real-world applications, the BCI's performance can be evaluated simply by determining the accuracy with which the user's intended commands are decoded by the system in real-time. However, in the preliminary stages of research, it is often more practical to collect and store a complete dataset that can be analyzed "offline" (i.e., not in real-time). This approach allows, for example, the comparison of different analysis methods. To evaluate the performance of the BCI algorithm when performing offline analysis, the data must be divided into training and test sets. The classifier is optimized based on the training set only, and its performance is evaluated on the unseen test set.

A single division into one training and one test set is a valid approach, and there are cases where it may even be necessary, for example, when it is important to preserve the order in which the data was collected. However, to reduce the variance resulting from the random division into data subsets and produce a more generalizable estimate of the classifier's performance, a cross-fold validation approach is typically followed.

K-fold cross-validation is a performance evaluation approach through which the training data is randomly divided into k distinct subsets (e.g., k=5 for an 80%/20% train/test split). A K-1 of these subsets/folds is then used to train the model, and the performance of the model is investigated using the remaining subset. This process is repeated until each subset is used as the test set once. The classifier performance is then estimated as the average of the classification accuracy obtained in each "fold" of the cross-validation algorithm. As a result, a more accurate estimation of the model's generalization is achieved. K-fold cross-validation is also often used in tuning classifier hyperparameters.

Chapter 3

Study 1: Investigating singing imagery as a potential control task for active BCI

A version of this chapter has been submitted as part of a manuscript titled "Investigating Singing Imagery as a Potential Control Task for Motor Imagery BCI", which is currently under review at IEEE Transactions on Biomedical Engineering (Manuscript ID: TBME-00207-2023).

In this chapter, we investigated the potential effectiveness of singing imagery as a control task for active EEG-based BCI systems, being careful to emphasize the speech and motor imagery aspects of the task. Considering the MI tasks that are most commonly used in EEG-BCI research (i.e., R, L, F, T, rest), we incorporated singing imagery into various binary and multi-class scenarios. The objective was to investigate the potential of singing imagery to 1) provide an alternative to the conventional motor imagery tasks in 2- to 5-class paradigms, and 2) to extend the number of possible commands from five (i.e., L, R, F, T, rest) to six (i.e., L, R, F, T, rest, SI).

3.1 Methods

3.1.1 Participants

15 healthy participants (mean age: 25±5.8; 12 right-handed; 10 Female) were recruited for this study. Participants were included if they had normal or corrected-to-normal vision and hearing, no cognitive impairment, and no history of neurological disease, disorder, or injury. Data recorded from one participant who reported a lack of engagement and was visibly drowsy during the experiment was excluded from analyses. Participants were asked to avoid exercising and consuming alcohol or caffeine for at least four hours before the experiment. All participants provided written informed consent prior to completing the experiment. The experimental protocol was approved by the Interdisciplinary Committee for Ethics in Human Research (ICEHR) at Memorial University of Newfoundland.

3.1.2 Data acquisition

For this experiment, EEG data was recorded via a 64-channel actiCHamp system with active electrodes (Brain Products, GmbH), sampling at 500 Hz. The impedance for all electrodes was kept below 10 KOhm throughout the experimental session. Electrodes were placed according to the international 10-10 system.

3.1.3 Experimental protocol

Each participant completed a single experimental session. The experimental protocol, summarized in Figure 3.1, is described in detail below.

The participants first completed the short version of the Kinesthetic and Visual Imagery Questionnaire (KVIQ-10) [40]. This is a standardized set of tasks that is commonly used in motor imagery BCI studies to subjectively evaluate the ability of participants to visualize and sense imagined movements. The KVIQ-10 ratings were collected to aid in interpreting individual participants' classification results.

Participants were seated comfortably in front of a standard computer monitor

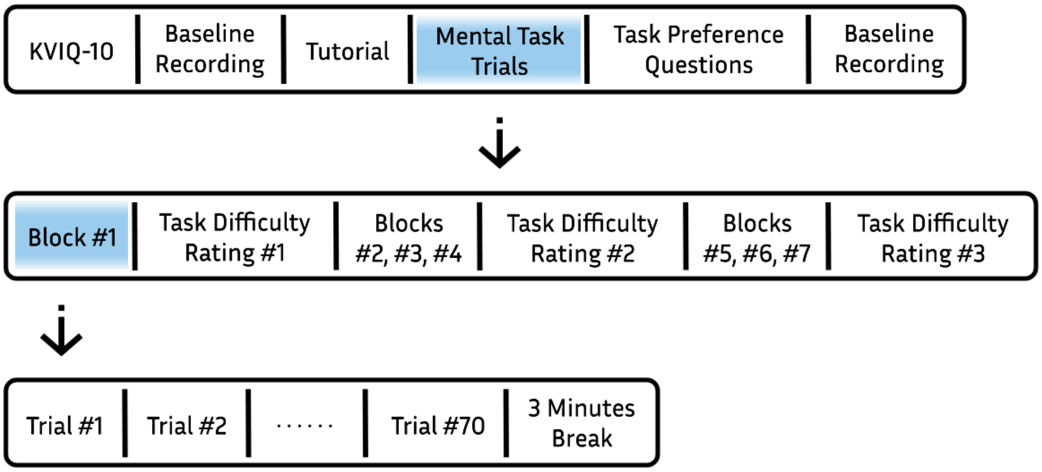

Figure 3.1: Experimental protocol illustrating the different steps of the single recording session. All of the blocks had a similar structure to block #1.

throughout the remainder of the experiment. Following the setup of the EEG electrodes, two one-minute baseline trials were recorded, one with eyes open and the other with eyes closed (these baseline trials were repeated at the end of the experiment as well). Before starting the motor imagery tasks, a description of the experimental protocol and imagery tasks was presented.

Next, participants completed the mental task trials. Seven different mental tasks (including rest) were performed. These tasks are described in detail below in section 3.1.3 (i.e., "Mental Tasks" section). Participants completed 70 trials per task (490 trials in total), which were completed in seven blocks of 70 trials (10 trials per task per block). Participants rested for a minimum of 3 minutes between blocks.

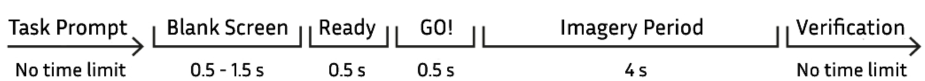

Figure 3.2: Timing of the individual trial.

Figure 3.2 illustrates the timing for each task trial. Each trial was 5.5-6.5 seconds in duration. At the start of each trial, the task to be performed in that trial was indicated on the screen. The participant began the timed portion of the trial when they felt ready to do so by pressing any key on the keyboard. Once a key was pressed, a blank screen appeared (0.5 – 1.5 seconds), followed by a "Ready!" (0.5 seconds) and then a "GO!" (0.5 seconds) screen, followed then by a dark screen with only a

"+" sign which indicated the period during which the participant was to perform the indicated mental task (4 seconds). The last step in each trial was the verification step.

The primary purpose of the verification step was to record the participant's perspective on the quality of their performance during the trial. Specifically, they chose one of the following three options to indicate their feeling about their performance of the task: 1) I performed it correctly, 2) I did not perform it correctly, and 3) I performed it correctly but not very well. Participants were instructed to choose "2) I did not perform it correctly" if they either did not perform the task at all, performed the wrong task, or moved during the task period. They were instructed to choose "3) I performed it correctly but not very well" if they performed the correct task but felt they did not perform it effectively (e.g., they were not well-focused or engaged in the task).

Mental tasks

Along with the conventional motor imagery tasks of the right hand (R), left hand (L), feet (F), and tongue (T), a rest state (REST) and two different singing imagery tasks ($\mathrm{SI}_{\mathrm{Kin}}$ and $\mathrm{SI}_{\mathrm{noKin}}$) were included. These tasks were defined as follows:

- **Right hand motor imagery (R):** Participants imagined tapping their right hand, bending at the wrist, at a steady pace of approximately one tap per second.
- **Left hand motor imagery (L):** Participants imagined tapping their left hand, bending at the wrist, at a steady pace of approximately one tap per second.
- **Feet motor imagery (F):** Participants imagined tapping both feet together, bending at the ankle, at a steady pace of approximately one tap per second.
- **Tongue motor imagery (T):** Participants imagined protruding their tongue out of their mouth and retracting it at a steady pace of approximately once per second.
- **Kinesthetic singing imagery ($\mathbf{SI}_{\mathbf{Kin}}$):** Participants imagined singing a song (with lyrics) in their heads, without vocalization or movement. In doing so, participants were instructed to focus on the kinesthetic sensation of moving their jaw, tongue, and lips as they imagined articulating the lyrics.

- **Non-Kinesthetic singing imagery ($\mathbf{SI_{noKin}}$):** Participants imagined singing a song (with lyrics) in their heads, without vocalization or movement, and without focusing on the kinesthetic sensations in the jaw, tongue and lips.
- **Rest (REST):** Participants were instructed that they do not need to do anything specific for this task. They were asked to keep their eyes open and focused on the screen (they could blink normally) and not perform any of the other six tasks in this experiment.

As previously mentioned, we view singing imagery as being either a special case of speech imagery or just a more intuitive and natural motor imagery task involving the jaw, tongue, and lips (or a combination of both). This is why we included two different singing imagery tasks, to emphasize these aspects of singing imagery and investigate if either version is more promising as a BCI control task. For both singing imagery tasks, participants selected the songs to imagine from a list of widely known English songs (e.g., Jingle Bells, Happy Birthday To You, The Alphabet Song). A different song was chosen for each block of trials.

Subjective evaluation of tasks

To better understand the participants' opinions on the different tasks, we asked them to rate the difficulty of each task after completing blocks 1, 4, and 7. They rated the tasks on a scale of one to five, one being not difficult at all and five being extremely difficult.

Furthermore, at the end of the session, they answered three multiple-choice questions about their task preferences. These questions were as follows:

1. Which four tasks would you pick to use on a daily basis to work with a computer?
    (a) R
    (b) L
    (c) F
    (d) T
    (e) SI_{Kin}
    (f) SI_{noKin}

2. Which task would you prefer to use in a BCI on a daily basis?

   (a) T

   (b) SI (Kin or noKin)

3. Which singing task did you find more intuitive and easier to perform?

   (a) SI_{noKin}

   (b) SI_{Kin}

3.1.4 EEG pre-processing

Preprocessing was conducted using the EEGLAB toolbox in MATLAB. This step consisted of down-sampling the data to 250 Hz, re-referencing to the average of all channels, epoch extraction, and baseline correction (i.e., removal of the mean for each epoch). Epochs were extracted for the period of -500ms until 4000ms from the onset of the timed portion of the trial.

3.1.5 Feature extraction

The Filter Bank Common Spatial Patterns (FBCSP) algorithm is one of the most efficient algorithms to extract spatio-spectral features from EEG signals. This algorithm has remained one of the most common and useful feature extraction techniques in MI-BCI studies since it was proposed [2]. In this study, we applied the FBCSP algorithm as described below:

Following epoch extraction, nine symmetric linear-phase FIR bandpass filters were used to filter the preprocessed EEG signals into bands of width 4 Hz, ranging from 4 Hz to 40 Hz (i.e., nine frequency bands in total). Next, for each frequency band, the Common Spatial Pattern (CSP) algorithm was used to fit optimal spatial filters to the EEG trial data such that the power of the resulting spatially filtered signals was maximally discriminant between the different populations (i.e., the different mental tasks being classified). In this study, using MNE library in python, 10 CSP filters

(5 pairs) were optimized for each of the nine frequency bands, yielding a total of 90 features.

For all classification problems considered, the Maximum Relevance Minimum Redundancy (mRMR) feature selection algorithm [50] was used to reduce the feature set dimensionality to 25. This value was selected based on preliminary analysis of the 6-class classification problems, where feature sets of size 6, 12, 20, 25, 45, 70, and 90 were investigated, and gains in accuracy appeared to plateau after about 25 features (see Fig 3.3).

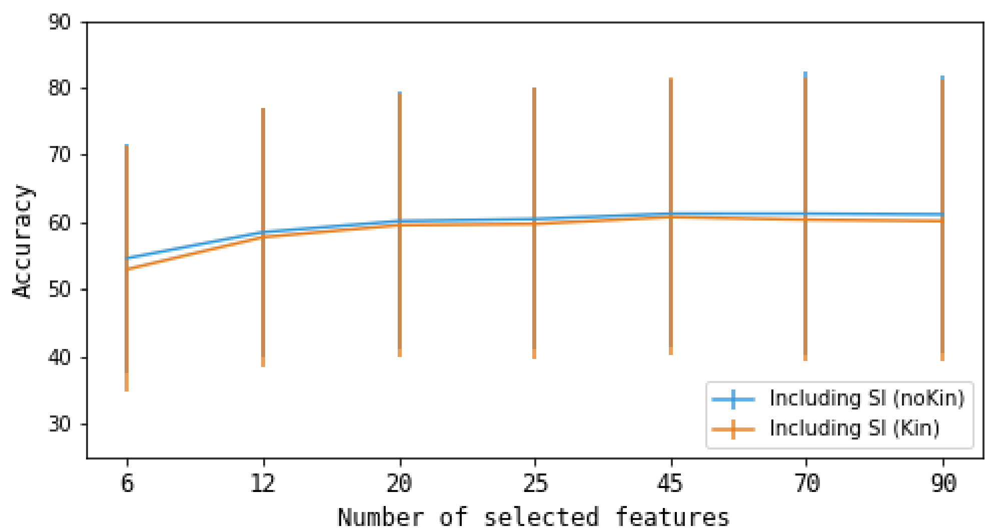

Figure 3.3: Effect of increasing number of selected features on the mean and standard deviation of accuracies for the 6-class scenarios.

3.1.6 Task classification via EEG

The objective of this work was to investigate the potential of singing imagery to provide an alternative to the conventional motor imagery tasks in 2- to 5-class paradigms and also to extend the number of possible commands from five (i.e., L, R, F, T, REST) to six (i.e., L, R, F, T, REST, SI). To do so, all possible combinations of the seven tasks (i.e., L, R, F, T, REST, SI_{noKin}, SI_{Kin}) were investigated for the 2-, 3-, 4-. 5- and 6-class scenarios (note that beyond the 2-class scenarios, combinations including both SI_{noKin} and SI_{Kin} were excluded). For each classification problem investigated, the average of five runs of 10-fold cross-validation was calculated to estimate the classifier's performance. A random forest classifier served as the learning algorithm as it is a powerful classification algorithm that has demonstrated its robustness in differentiating EEG signal populations [43]. In each fold of the cross-validation algorithm, no test data was involved in feature extraction, feature selection, or classifier optimization.

3.1.7 Self-evaluation data

Effect of discarding potentially "bad" data on classification accuracy

We thought that the self-evaluation conducted by the participants at the end of each trial could potentially indicate "bad" data that can/should be discarded. Based on the responses to the verification step of each trial, the 6-class problems were repeated under two scenarios:

1) With all trials for which the participants responded "I did not perform it correctly" or "I performed it correctly but not very well" removed.

2) With all trials for which the participants responded "I did not perform it correctly" removed.

The results in each of these conditions were compared to the results obtained when equivalent numbers of trials were randomly removed from the dataset (to ensure that the size of the training sets are comparable).

Correlation analyses

- Self-evaluation and Classification Accuracy

The correlation between the results of participants' self-evaluation and the classification accuracies achieved in the 6-class scenarios (with no trials discarded) was investigated. Specifically, the correlation was calculated between the 6-class classification accuracy and the percentage of trials for which the participants responded 1) "I performed it correctly", and 2) "I performed it correctly" or "I performed it correctly but not very well".

- KVIQ-10 Scores and Task Differentiability

The correlation between the participants' ability to perform motor imagery, as determined by their KVIQ-10 scores, and the classification accuracies achieved in the 6-class scenarios (with no trials discarded) was investigated. Specifically, the correlation was calculated between the 6-class classification accuracy and the participants' 1) visual imagery, 2) kinesthetic imagery, 3) overall KVIQ-10 scores.

- Self-evaluation and KVIQ-10 Scores

An investigation of a potential relationship between the participants' KVIQ-10 scores and their subjective task performance verification was conducted. Specifically, the correlation was calculated between the KVIQ-10 scores and the percentage of trials for which the participants responded 1) "I performed it correctly", and 2) "I performed it correctly" or "I performed it correctly but not very well".

3.2 Results

3.2.1 Participant's tasks preferences and difficulty ratings

Figure 3.4. summarizes the participants' responses to the question *"Which four tasks would you pick to use on a daily basis to work with a computer?* For each task, the number of participants who included it as one of their top 4 choices is indicated. In response to the other two questions regarding task preferences, 13 out of 14 participants indicated that they preferred the singing imagery tasks to the tongue motor imagery task, and 8 out of 14 participants indicated that they preferred SI_{Kin} to SI_{noKin}.

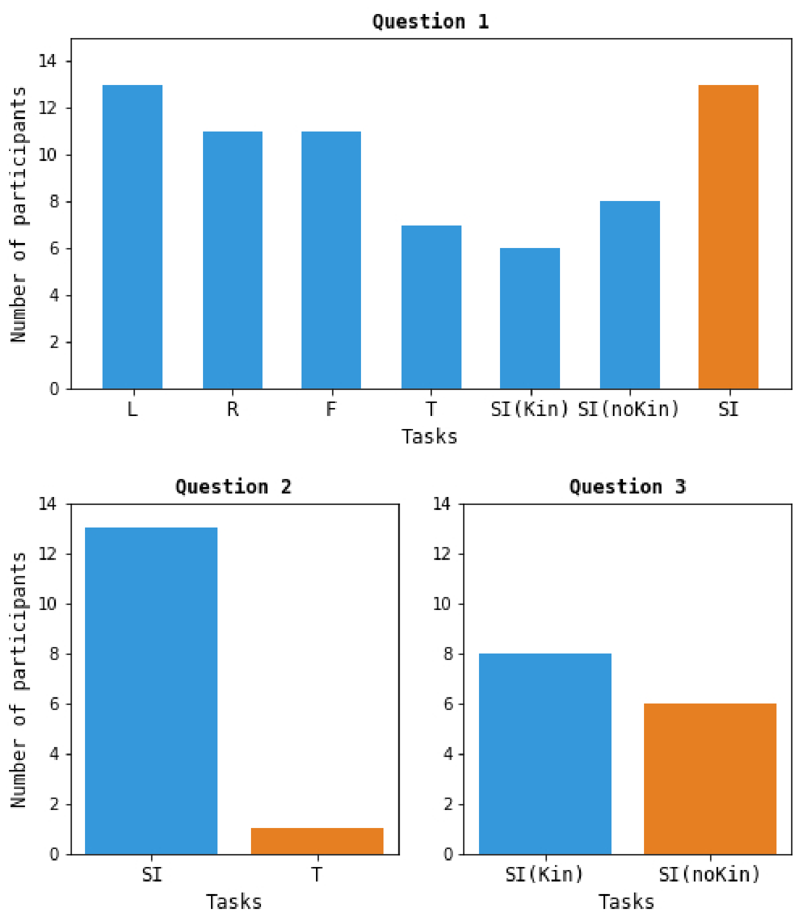

Figure 3.4: Participant responses to the task preference questions. The "SI" bar in question 1's chart indicates the number of participants who included either one of the two SI tasks as one of their top 4 choices.

Participants were asked to rate the difficulty of each task on a five point scale three times during the experiment - after completing Blocks 1, 4 and 7. The average difficulty ratings (across participants) are reported in Table 3.1.

Table 3.1: Task difficulty ratings

	Recording after			**Task Average**
	Block 1	**Block 4**	**Block 7**	
L	2.1 (1.0)	1.9 (0.7)	1.9 (0.9)	**2.0 (0.8)**
R	2.2 (0.8)	2.0 (0.8)	1.9 (0.9)	**2.1 (0.7)**
F	1.9 (0.8)	2.0 (1.2)	1.9 (1.1)	**2.0 (1.0)**
T	2.9 (1.4)	2.2 (1.1)	2.2 (1.2)	**2.4 (1.1)**
$\mathbf{SI_{Kin}}$	2.6 (1.5)	2.4 (1.2)	2.4 (1.4)	**2.5 (1.3)**
$\mathbf{SI_{noKin}}$	2.2 (1.3)	1.9 (0.8)	1.8 (0.8)	**2.0 (1.0)**
REST	1.6 (1.0)	1.4 (0.7)	1.2 (0.4)	**1.4 (0.9)**
Block Average	**2.2 (0.6)**	**2.0 (0.6)**	**1.9 (0.6)**	

Block average is the average perceived difficulty of all the tasks after that block and task average is the average perceived difficulty of that specific task recorded at three different points during the experiment.

3.2.2 EEG classification

The results for the different classification scenarios (i.e., binary, 3-class, 4-class, 5-class, and 6-class) are reported in Tables 3.2, 3.3, 3.4, 3.5, and 3.6, respectively. The reported values are the grand average classification accuracies and the calculated standard deviation (in parenthesis) across all the participants. For each scenario, results for all combinations of the seven tasks are included (except those combinations that include both types of SI, which is only reported for the binary scenario). Also, in the 2-class scenarios, the accuracy of classifying SI_{Kin} against SI_{noKin} was not included in the calculation of average accuracy for these tasks.

Furthermore, the per participant classification accuracies for the 6-class scenarios are shown in Figure 3.5. The 6-class analysis was repeated with randomized class labels to produce a distribution for each participant representing "chance" level classification; paired t-tests revealed that the 6-class accuracies displayed in Figure 3.5

Table 3.2: Classification accuracies for the 2-class scenarios (%)

2-Class scenarios (chance = 50%)							
	Conventional Motor Tasks			**Singing Imagery**		**REST**	**Average**
	R	**F**	**T**	**SI_{Kin}**	**SI_{noKin}**		
L	81.7 (15.0)	86.7 (13.7)	89.9 (9.5)	90.0 (9.0)	90.5 (8.3)	87.5 (10.4)	87.7 (3.3)
R	-	87.2 (13.3)	89.2 (9.5)	88.7 (10.2)	88.2 (9.3)	85.6 (10.9)	86.8 (2.8)
F	-	-	82.6 (14.2)	84.5 (12.7)	85.9 (11.3)	83.6 (12.1)	85.1 (1.9)
T	-	-	-	71.8 (12.8)	78.9 (13.5)	82.7 (11.2)	82.5 (6.7)
SI_{Kin}	-	-	-	-	67.6 (14.2)	83.1 (12.9)	83.5 (7.2)
SI_{noKin}	-	-	-	-	-	74.1 (12.1)	83.6 (6.8)
REST	-	-	-	-	-	-	82.8 (4.6)

Table 3.3: Classification accuracies for the 3-class scenarios (%)

3-Class scenarios (chance = 33.3%)					
	Conventional Motor Tasks		**Singing Imagery**		**REST**
	F	**T**	**SI_{Kin}**	**SI_{noKin}**	
L + R	78.3 (18.0)	80.9 (14.0)	80.2 (15.9)	79.3 (16.4)	78.5 (15.9)
L + F	-	76.4 (16.5)	79.3 (15.8)	79.1 (15.8)	77.9 (16.1)
R + F	-	78.2 (15.9)	78.8 (15.6)	78.2 (15.4)	77.1 (16.1)
L + T	-	-	73.0 (13.3)	77.9 (13.8)	77.6 (14.7)
R + T	-	-	72.5 (13.0)	76.7 (14.4)	77.7 (13.7)
T + F	-	-	66.7 (16.1)	70.8 (17.1)	71.4 (17.2)
L + REST	-	-	78.0 (14.4)	74.1 (11.2)	-
R + REST	-	-	77.5 (14.0)	72.5 (11.9)	-
F + REST	-	-	73.6 (16.8)	69.9 (13.8)	-
T + REST	-	-	65.6 (15.2)	65.8 (13.4)	-

Table 3.4: Classification accuracies for the 4-class scenarios (%)

4-Class scenarios (chance = 25%)			
L + R + F + T	72.8 (18.3)		
	Singing Imagery		**REST**
	SI_{Kin}	**SI_{noKin}**	
L + R + F	73.7 (18.4)	73.7 (19.0)	72.6 (19.0)
R + F + T	67.6 (17.1)	70.0 (18.0)	70.0 (17.7)
L + F + T	66.6 (16.8)	70.2 (18.1)	69.5 (19.0)
R + L + T	69.0 (16.5)	73.3 (17.0)	72.5 (17.8)
F + T + REST	62.0 (17.3)	63.1 (15.8)	-
R + T + REST	66.0 (15.4)	67.5 (14.1)	-
R + F + REST	70.6 (18.0)	68.3 (15.5)	-
L + T + REST	67.0 (15.5)	68.0 (14.6)	-
L + F + REST	71.4 (17.9)	68.1 (15.4)	-
R + L + REST	71.9 (18.0)	69.8 (15.6)	-

Table 3.5: Classification accuracies for the 5-class scenarios (%)

5-class scenarios (chance = 20%)		
L + R + F + T + REST	66.5 (20.1)	
	Singing Imagery	
	SI_{Kin}	**SI_{noKin}**
L + R + F + T	63.8 (18.8)	67.1 (19.8)
L + F + T + REST	62.1 (18.3)	62.1 (18.0)
R + F + T + REST	62.9 (17.6)	62.1 (18.0)
L + R + T + REST	63.6 (17.7)	64.6 (17.7)
L + R + F + REST	66.9 (20.3)	65.5 (18.0)

Table 3.6: Classification accuracies for the 6-class scenarios (%)

6-Class scenarios (chance = 16.7%)		
	Singing Imagery	
	SI_{Kin}	**SI_{noKin}**
L + R + F + T + REST	59.7 (19.7)	60.7 (19.2)

are significantly greater than chance for all participants (t>9.27; p<0.001). It should be noted that one of the participants (P8) finished only four blocks of mental tasks.

To indicate how well each class is being predicted in the 6-class scenario, Figure 3.7 shows the average confusion matrices for the two possible 6-class combinations (i.e., five conventional tasks + SI_{Kin} , and five conventional tasks + SI_{noKin}).

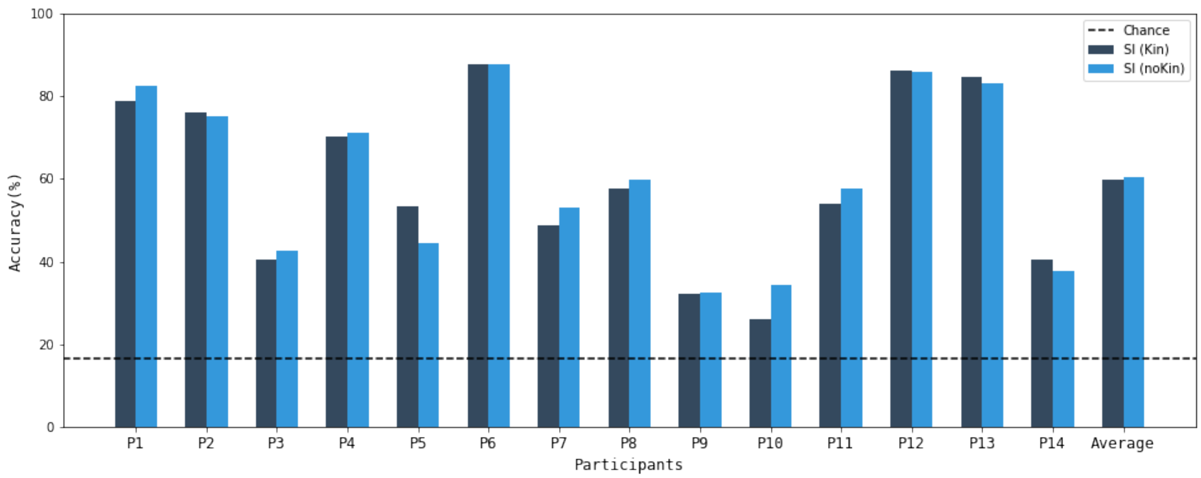

Figure 3.5: 6-class classification accuracies achieved for all the participants in this study.

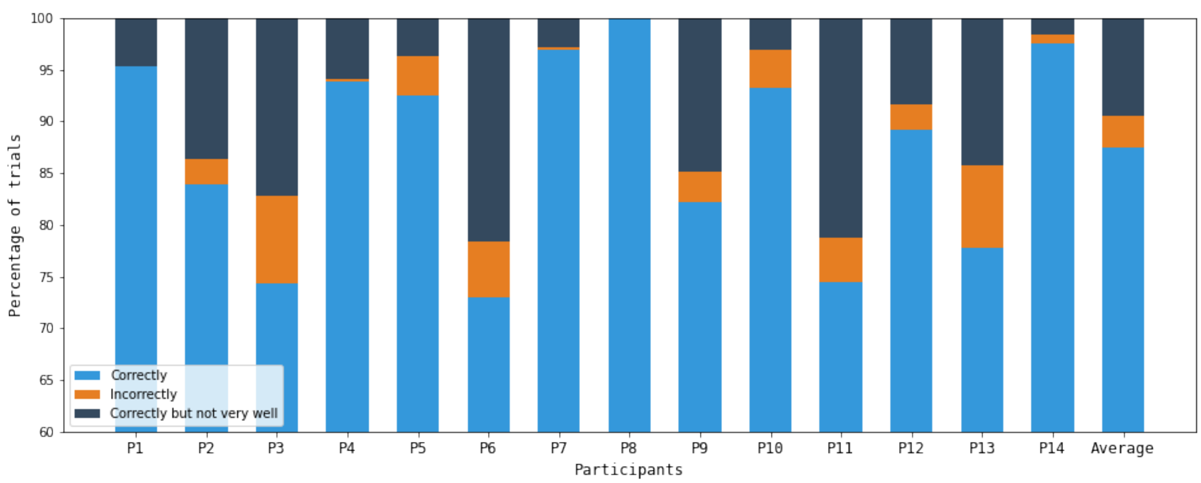

Figure 3.6: Results of self-evaluation for all of the participants.

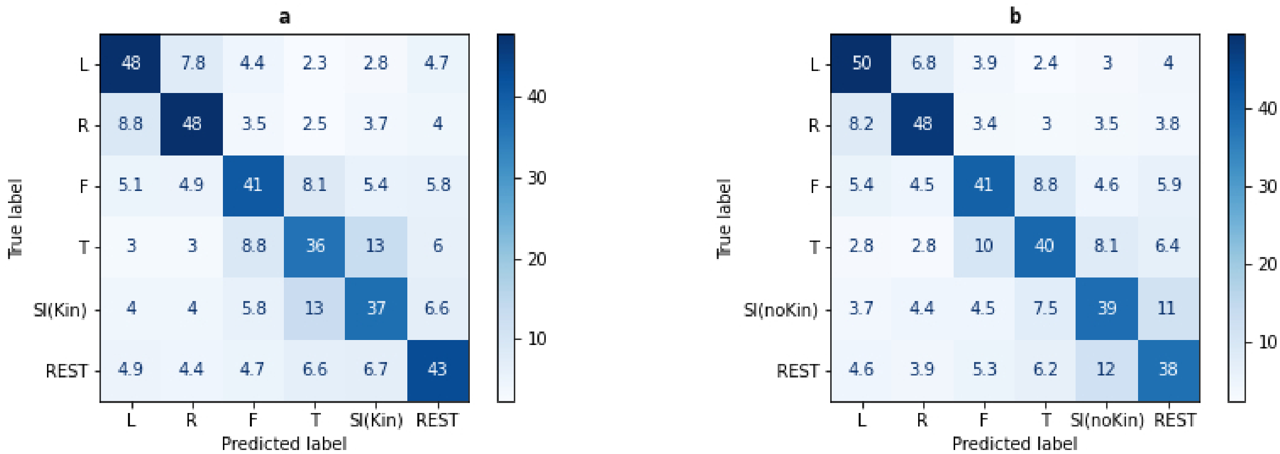

Figure 3.7: Confusion matrices for the two 6-class scenarios. a) Combination with SI_{Kin} b) Combination with SI_{noKin}. Results of P8, which only had 4 blocks of data (and, as a result, a different total number of trials), were excluded in the average confusion matrices.

Table 3.7 shows the average of precision, recall, and F1 metrics for the two 6-class combinations. These metrics are defined in equations 3.1, 3.2, and 3.3.

$$Precision = \frac{TP}{TP + FP} \tag{3.1}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.2}$$

$$F1score = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \tag{3.3}$$

where

TP: Number of times a task was correctly labeled as being a member of a specific class

FP: Number of times a task was incorrectly labeled as being a member of a specific class

FN: Number of times a member of a specific class was wrongly labeled as being a member of another class.

Table 3.7: Precision, Recall, and F1 metrics calculated for different tasks in the 6-class scenarios (%).

Performance evaluation of tasks in 6-class scenarios							
		Tasks					
SI type	**Metric**	**L**	**R**	**F**	**T**	**SI**	**REST**
$\mathbf{SI_{Kin}}$	Precision	67.2 (26.3)	66.4 (23.4)	58.6 (25.6)	51.8 (17.7)	52.0 (17.2)	60.0 (20.9)
	Recall	69.6 (25.1)	68.7 (24.7)	59.3 (28.8)	50.9 (18.0)	50.8 (17.9)	59.0 (21.1)
	F1	68.3 (25.7)	67.4 (24.0)	58.7 (27.2)	51.3 (17.8)	51.3 (17.5)	59.3 (20.9)
$\mathbf{SI_{noKin}}$	Precision	68.9 (24.1)	68.3 (21.4)	59.0 (24.8)	58.3 (22.9)	53.6 (17.2)	54.1 (15.8)
	Recall	71.8 (22.5)	69.7 (22.7)	59.3 (27.0)	56.4 (23.3)	53.8 (19.2)	52.9 (15.2)
	F1	70.3 (23.3)	68.9 (21.9)	58.9 (25.8)	57.3 (23.1)	53.5 (18.1)	53.4 (15.5)

3.2.3 Self-evaluation data

The results of self-evaluation for each participant are summarized in Fig 3.6. For each participant, the percentage of trials which they indicated were performed “correctly”, “incorrectly” or “correctly but not very well” are reported.

Effect of discarding potentially “bad” data on classification accuracy

Table 3.8 shows the results of the 6-class classification analysis when potentially “bad” trials, as indicated by the participants at the verification step of each trial, were removed. Accuracies reported are grand average balanced accuracies across all participants. Removing the trials for which participants responded either “I did not perform it correctly” or “I performed it correctly but not very well” did result in slight, but not significant, enhancement of classification accuracy (around 1.2%).

Table 3.8: Discarding potentially “bad” data in the 6-class scenarios (%)

6-Class scenarios with discarding potentially “bad” data				
	Scenario #1		**Scenario #2**	
SI Type	**All “Incorrectly” and “Correctly but not very well” trials discarded**	**Same number of trials randomly discarded**	**All “Incorrectly” trials discarded**	**Same number of trials randomly discarded**
$\mathbf{SI_{Kin}}$	60.3 (20.1)	59.3 (20.1)	60.9 (19.8)	59.7 (20.0)
$\mathbf{SI_{noKin}}$	60.8 (19.7)	59.6 (19.9)	61.5 (19.3)	60.3 (19.7)

Correlation analyses

- Self-evaluation and Classification Accuracy

The results did not reveal any significant correlation between how the participants judged their performance and the ability of the classifier to differentiate the tasks ($p>0.05$).

- KVIQ-10 Scores and Task Differentiability

The KVIQ-10 ratings also appear to have no significant correlation with the achieved accuracies ($p>0.05$).

- Self-evaluation and KVIQ-10 Scores

The results of self-evaluation do not demonstrate any significant correlation with the KVIQ-10 ratings ($p>0.05$).

3.3 Discussions

3.3.1 Task preferences and difficulty ratings

The difficulty of the tasks used to control a BCI can significantly impact its practicality and user-friendliness. For an optimal design, one should consider not only how accurately the task can be detected by the classifier but also how easy it is to perform the task from the user's perspective. Given that the target users of active BCI systems are typically individuals with severe physical disabilities, the importance of incorporating tasks that feel natural and do not require long training sessions is even greater.

Evaluating the results of the survey which asked the participants about their top four preferred tasks, the majority of participants seemed to prefer motor imagery of hands and feet, while the motor imagery of tongue and the individual SI tasks were chosen by fewer participants. However, it should be noted that in the case of SI tasks, 13 out 14 participants included either one of the two SI tasks as one of their top four preferred tasks. Moreover, based on the results of question 2, 13 out of 14 participants found the singing imagery tasks preferable to tongue imagery. In terms of the two singing imagery tasks, there seemed to be a relatively even split between participants who preferred SI_{Kin} (8) and SI_{noKin} (6). The difficulty ratings, for the most part, paint a similar picture of the participants' perception of the tasks. Unsurprisingly participants found the REST task the easiest task in the list. Following that, motor imagery of R, L, F, and SI_{noKin} all had similar difficulty ratings (i.e., around 2 out of 5), while motor imagery of T and SI_{Kin} were more challenging for the participants (i.e., around 2.5 out of 5). One apparent contradiction in the results is that despite SI_{Kin} being rated as more difficult than SI_{noKin} on average, more participants indicated that they preferred it.

At first look, the recorded task difficulty ratings suggest that participants did not find performing the singing imagery straightforward. However, based on informal feedback we received from the participants, distinguishing the instructions for the two types of singing imagery may have increased their perceived difficulty. Specifically, some of the participants reported that they rated one or both of the singing imagery tasks as difficult just because they struggled to follow the specific instructions of focusing on the kinesthetic feeling of movement in one case and avoiding that for the other scenario. Therefore, it can be assumed that this difficulty would be significantly relieved in a final design where only one clear instruction for the SI tasks is included.

3.3.2 EEG classification results

2-Class scenarios

- Task vs. REST Classification

To reach a practical design, any new task should be differentiable from the rest state. In this study, R and L yielded the highest classification accuracy against the rest state (85.6% and 87.5%, respectively). Encouragingly, accuracy for SI_{Kin} vs. REST (83.1%) proved to be comparable to F vs. REST and T vs. REST (83.6% and 82.7%, respectively). However, the classification accuracy obtained for SI_{noKin} vs. REST was considerably lower (74.1%). This may indicate that focusing on the kinesthetic feeling of the movement in the jaw, tongue, and lips during singing imagery generates stronger and/or more distinct patterns of brain activity and, therefore, may be a more effective version of the task for BCI control.

- Task vs. Task Classification

Before including the SI tasks in multi-class scenarios and to investigate their potential utility in binary systems, it is important to see how differentiable they are from the other tasks.

Based on the results obtained here, both SI tasks appear to be equally useful in classifying against the conventional motor imagery tasks. The only notable difference between the two SI tasks appears to be in classifying against the tongue imagery task where SI_{noKin} outperformed SI_{Kin} by 7% (78.9% as compared to 71.8%). Because for the SI_{Kin} task, participants were explicitly asked to focus on the sensations involved with moving the jaws, lips, and tongue, it makes sense that similar brain areas would be activated as in the tongue imagery task, and the task differentiability would be reduced. Unsurprisingly, the classification of SI_{noKin} vs SI_{Kin} yielded the lowest classification accuracy among all binary problems, with an average accuracy of 67.6%. Overall, the classification accuracies for the remaining task vs. task problems were more or less comparable and fell in the 80-90% range. The highest accuracies were for the R or L vs. either T, SI_{noKin} or SI_{Kin} (these accuracies all ranged between 88.2-90.5%). The lowest accuracies were obtained for L vs. R (81.7%) and F vs. T (82.6%). In comparison to T, SI_{Kin} and SI_{noKin} both provide slightly higher classification (2-3%) accuracies against F.

3 , 4, and 5-Class scenarios

The results for the multi-class scenarios verify the potential of singing imagery as an alternative for different motor tasks. Specifically, by incorporating singing imagery, accuracies as high as 80.2%, 73.7%, and 67.1% were achieved for 3, 4, and 5-class combinations, respectively. The conventional combinations for the 4, and 5-class problems (i.e., R + L + F + T and R + L + F + T + REST, respectively) resulted in average accuracies of 72.8% and 66.5%. Therefore, based on these results, it is feasible to use singing imagery for developing 3, 4, and 5-class paradigms that provide accuracies that are comparable to combinations of conventional tasks and are potentially more user-friendly.

Based on our finding from the 2-class scenarios, some pairs of tasks appeared to generate similar patterns of activity and were, therefore, harder to distinguish for the classifier (e.g., T and SI_{Kin}). Multi-class scenarios appear to follow similar trends where including both such tasks in the combination results in a decline of performance in comparison to combinations that only use one of them. (F + T), (T + SI_{Kin}), and (REST + SI_{noKin}) are pairs of tasks that, when used together in multi-class scenarios, lead to accuracies as low as 65.6%, 62.0%, and 62.1% for the 3, 4, and 5-class scenarios, respectively. Interestingly, the minimum accuracy for 4 and 5 class combinations are comparable.

Having the BCI be able to recognize a "No-control" or "Rest" state is crucial for a more practical and user-friendly design as it means that users would not be required to perform a task during all control periods. All the multi-class scenarios involving the SI and REST task achieved accuracies well above the chance level. For the 3, 4, and 5-class scenarios, accuracies for such combinations were in the ranges of 65.6-78.0%, 62.0-71.9%, and 62.1-66.9%, respectively.

6-class problems

The potential of increasing the number of possible commands in MI-BCI through incorporating singing imagery is supported through the results of the 6-class analysis. Average accuracies of around 60% were obtained, which is more than three times the chance level (i.e., 16.7%). While the results did not, on average, reach 70% (the accuracy often cited as being necessary for effective communication [34]), they are not far off, and may be obtainable with further optimized pre-processing and classification techniques or potentially with training/practice by the users. Still, it is encouraging that accuracies exceeding 70% were achieved for 6 out of 14 participants, and all participants' accuracies were significantly higher than chance. Both of the singing imagery tasks yielded comparable results, with neither significantly outperforming the other.

The confusion matrices obtained from 6-class scenarios further support the idea mentioned above that the motor imagery of tongue and singing imagery (especially SI_{Kin}) generate similar patterns of brain activity. The most common error for the classifier on the combination that involves SI_{Kin} is in labeling T as SI_{Kin} or SI_{Kin} as T. On the other hand, for the combination that involved SI_{noKin}, differentiating REST and SI_{noKin} proved to be challenging for the classifier. Furthermore, as the previous results also suggest, the motor imagery of F seems to be more distinct from singing imagery than it is from the motor imagery of T. Thus again, these results appear to be aligned with the results of binary classification.

3.3.3 Self-evaluation data

The results of the 6-class classification when potentially "bad" trials were removed, along with the lack of correlation between the classification accuracy and the verification responses, suggests that the participants' perception of how good or bad they performed the tasks may be unreliable. Similarly, KVIQ-10 scores –visual, kinesthetic, and total- did not demonstrate any significant correlation with the 6-class classification accuracies. This is in contradiction with the findings of Vuckovic et al. [65], where KVIQ-10 scores were reported to be capable of predicting BCI illiteracy of participants. Moreover, since the participants themselves are reporting both the KVIQ-10 ratings and the task verification responses (i.e., correctly, incorrectly, or correctly but not very well), one may assume that there may be a correlation between how they perceive their performance throughout the trials and how they felt about each of the KVIQ-10 tasks. However, our findings do not support this idea, and these recordings were not correlated. Thus, none of the subjective feedback recordings had a significant correlation with the 6-class classification accuracies or with each other. This could potentially indicate that the subjective information is not a reliable predictor for the capabilities of a participant in performing imagery tasks.

3.4 Conclusion

Our investigations indicate that singing imagery could provide a robust and potentially more intuitive alternative for conventional motor tasks in designing active BCIs. Moreover, the results suggest that it may feasible to use singing imagery for increasing the number of distinguishable commands to six and achieve average accuracies around 60%. An increase in the number of commands would lead to an enhancement in the practicality and functionality of conventional active BCI systems.

Chapter 4

Study 2: Investigating dual imagery of singing and motor imagery as potential control task for active BCI

A version of this chapter has been submitted as part of a manuscript titled "Investigating Dual Imagery Tasks for BCI Control", which is currently under review at the Journal of Neural Engineering (Manuscript ID: JNE-106478).

In this study, the potential effectiveness of DI tasks as control tasks for active EEG-based BCI systems was investigated. Considering the MI tasks that are most commonly used in EEG-BCI research (i.e., R, L, F), we combined each of them with singing imagery to create a novel set of mental tasks. Similar to Study 1, described in Chapter 3, one of the main objectives was to investigate the potential of DI tasks for increasing the number of commands for active BCIs.

4.1 Methods

4.1.1 Participants

14 healthy participants (mean age: 25.1±11.6; all right-handed; seven male) were recruited for this study. Participants were included if they had normal or corrected-to-normal vision and hearing, no cognitive impairment, and no history of neurological disease, disorder, or injury. Two of the participants did not finish the entire experimental protocol and performed only two blocks of mental tasks. Hence, these participants were excluded from the analyses and the reported results. Participants were asked to avoid exercising and consuming alcohol or caffeine for at least four hours before the experiment. All participants provided written informed consent prior to completing the experiment. The experimental protocol was approved by the Interdisciplinary Committee for Ethics in Human Research (ICEHR) at Memorial University of Newfoundland.

4.1.2 Data acquisition

For this experiment, EEG data was recorded via a 64-channel actiCHamp system with active electrodes (Brain Products, GmbH), sampling at 500 Hz. The impedance for all electrodes was kept below 10 KOhm throughout the experimental session. Electrodes were placed according to the international 10-10 system.

4.1.3 Experimental protocol

Each participant completed a single experimental session during which they completed trials of each of the MI tasks of interest. The procedure of the experiment, which is summarized in Figure 4.1, is explained thoroughly in the following section.

The participants started by completing the Kinesthetic and Visual Imagery Questionnaire (KVIQ-10) [40]. This questionnaire is a standard set of tasks that are frequently used in BCI studies to subjectively assess the participants' ability to visualize and feel the sensations associated with kinesthetic imagined movements. The KVIQ-10 scores could potentially be helpful in interpreting the results of the MI trial data.

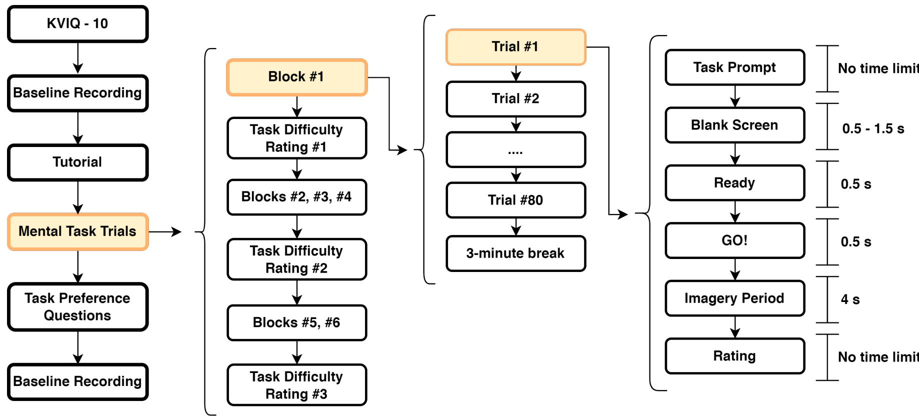

Figure 4.1: Experimental protocol illustrating the different steps of the single recording session. All of the blocks had similar structure to block #1.

The participant was seated facing a computer screen for the duration of the experiment. The chair and monitor's height were adjusted to the participant's comfort level. After the EEG electrodes were set up, two one-minute baseline trials, one with eyes open and the other with eyes closed, were recorded. These two baseline trials were repeated at the end of the experiment as well. Prior to beginning the motor imagery tasks, a brief tutorial explaining the experimental process and the imagery tasks was provided to the participant.

The participant then started performing the mental task trials. They were asked to perform eight different tasks, including rest. A detailed description of these tasks is given in section 4.1.3 (i.e., "Mental Tasks" section). Each task was performed 60 times for a total of 480 trials. Trials were completed in 6 blocks of 80 trials each (10 trials per task per block). The task order was random in each block. The participants were instructed to take a break of at least 3 minutes between blocks.

The timeline for each trial is also shown in Figure 4.1. At the beginning of each trial, the task to be performed was displayed on the screen. Participants were able to initiate the timed portion of the trial, which lasted between 5.5-6.5 seconds, by pressing any key on the keyboard. The screen then turned blank for a brief period (0.5 – 1.5 seconds), followed by the message "Ready!" (0.5 seconds), then "GO!" (0.5 seconds). After this, a dark screen with only a "+" sign appeared, signaling the time for the participant to perform the mental task (4 seconds). The final step of each trial was the performance rating step.

The performance rating step, repeated at the end of each trial, was an effort to try and capture the participant's perception of how well they performed the task. Specifically, they chose one of the following four options to indicate their feeling about their performance of the task: 1) Strong performance, 2) Moderate performance, 3) Weak performance, and 4) Wrong performance. Participants were asked to rate their performance based on factors such as the intensity of the sensations in their muscles, their focus during the task imagination period, and their general feeling about that trial. They were told to choose the option "4) Wrong performance" if they either didn't perform the task, did the wrong task, or moved during the imagery period.

Mental tasks

For this experiment, a singing imagery (SI) task, three single conventional motor imagery (MI) tasks, and three dual imagery (DI) tasks (i.e., the simultaneous performance of one of the MI tasks and SI) were included. A "rest" state was also included. The instructions for SI were based on Study 1, where a focus on the kinesthetic feeling of movement in the jaw and tongue appeared to be helpful in generating more distinct patterns of activity. These tasks were defined as follows:

- **Rest (REST):** Participants were instructed that they do not need to do anything specific for this task but were asked not to perform any of the other seven tasks in the experiment. They were asked to keep their eyes open and focused on the screen throughout the trial (they could blink normally).

- **Singing imagery (SI):** Participants imagined singing a song (with lyrics) in their heads, without vocalization or movement. In doing so, participants were instructed to focus on the kinesthetic sensation of moving their jaw, tongue, and lips as they imagined articulating the lyrics.

- **Left hand motor imagery (L):** Participants imagined tapping their left hand, bending at the wrist, at a steady pace of approximately one tap per second.

- **Right hand motor imagery (R):** Participants imagined tapping their right hand, bending at the wrist, at a steady pace of approximately one tap per second.

- **Feet motor imagery (F):** Participants imagined tapping both feet together, bending at the ankle, at a steady pace of approximately one tap per second.
- **Dual left hand imagery ($\mathbf{L_{SI}}$):** Participants imagined performing both the left hand and singing imagery tasks simultaneously. Specifically, participants were instructed to imagine singing the lyrics of the song and tapping their left hand to the rhythm of the song.
- **Dual right hand imagery ($\mathbf{R_{SI}}$):** Participants imagined performing both the right hand and singing imagery tasks simultaneously. Specifically, participants were instructed to imagine singing the lyrics of the song and tapping their right hand to the rhythm of the song.
- **Dual feet imagery ($\mathbf{F_{SI}}$):** Participants imagined performing both the feet and singing imagery tasks simultaneously. Specifically, participants were instructed to imagine singing the lyrics of the song and tapping their feet to the rhythm of the song.

For the DI tasks, participants were instructed to focus on feeling the kinesthetic sensations of movements for both tasks (i.e., to imagine the feeling of moving their jaw, tongue, and lips while singing, and their hand or feet while tapping). Similar to Study 1, for the singing imagery and DI tasks, participants selected the songs to imagine from a list of widely known English songs (e.g., Jingle Bells, Happy Birthday To You, The Alphabet Song). A different song was chosen for each block of trials.

Subjective evaluation of tasks

The participants were asked to express their opinion about the difficulty of each task after completing blocks 1, 4, and 6. The rating was based on a scale of 1 to 5, where 1 represented "not difficult at all" and 5 represented "extremely difficult."

Additionally, after finishing all the imagery trials, the participants were asked questions to help determine their task preferences. They were asked the question:

"If you were to choose one of the tasks below to do on a daily basis to use a computer, which one would you pick?"

They were presented with the following three scenarios and asked to select one of the two options for each:

1) **a)** Single tasks, **b)** Dual tasks

2) **a)** Hand imagery + Singing imagery, **b)** Feet imagery + Singing imagery

3) **a)** Right hand + Singing imagery, **b)** Left hand + Singing imagery

4.1.4 EEG pre-processing

The EEGLAB toolbox in MATLAB was used for pre-processing the data. The pre-processing pipeline involved several steps aimed at preparing the data for further analysis: removing the DC component, re-referencing the data to the average of all channels, applying an anti-aliasing filter, down-sampling the data to 250 Hz, extracting the epochs, and baseline correction (i.e., by removing the epoch's mean) were performed. The epochs were extracted for the time period ranging from -500ms to 4000ms, starting from the onset of the timed portion of each trial.

4.1.5 Feature extraction

The FBCSP algorithm was utilized to extract spatio-spectral features from the EEG signals. FBCSP is one of the most effective algorithms for this purpose and has been widely used in motor imagery BCI studies since its introduction [2]. Also, this algorithm successfully captured the features related to the singing imagery task in Study 1.

Because of the inherent similarities of the tasks in this study (i.e., the single tasks and the dual tasks), to fully capture their corresponding patterns, an extended version of this algorithm (as compared to that used in Study 1) was applied that extracted more detailed spectral information.

Following pre-processing, a set of symmetric linear-phase FIR bandpass filters were used to filter the pre-processed EEG signals into bands of width 4, 8, 12, and 16 Hz ranging from 0 Hz to 96 Hz. There was no frequency overlap for the 4 Hz filters, but for the rest of the filters, an overlap equal to half the filter's bandwidth was used. This resulted in a total of 73 frequency bands. Next, for each frequency band, the

Common Spatial Pattern (CSP) algorithm was used to transform the EEG data. The CSP implementation available on the MNE library was used to optimize 10 CSP filters (5 pairs) for each of the frequency bands, yielding a total of 730 features.

4.1.6 Task classification via EEG

The main objective is to evaluate the possibility of using DI tasks to increase the number of commands for active BCIs. The first step was to determine if it is possible to differentiate the DI tasks and each of the constituent single tasks. Thus the following 3-class scenarios were investigated: 1) R vs. SI vs. RSI, 2) L vs. SI vs. LSI, and 3) F vs. SI vs. FSI.

Next, incorporating the rest state into these problems, the following 4-class scenarios were investigated: 1) R vs. SI vs. R_{SI} vs. REST, 2) L vs. SI vs. L_{SI} vs. REST, and 3) F vs. SI vs. F_{SI} vs. REST.

Finally, the possibilities of increasing the number of commands were investigated through 7- and 8-class classification of the single and dual tasks. The 8-class scenario covered all the tasks performed in this experiment, and the 7-class combinations were obtained by removing the tasks, one at a time, from the 8-class combination.

A regularized LDA classifier, which shrinks the feature set and avoids overfitting or underfitting, was used for classification. The large size of the feature set could be partially addressed by using the linear discriminant analysis algorithm, as it is a fairly straightforward and computationally efficient algorithm. The classifier performance was calculated through five runs of 10-fold cross-validation. No test data was used in optimizing the CSP filters or the classifier.

4.1.7 Correlation between classification accuracy and self-rating data

For each subject, the Pearson correlation coefficient between the 8-class classification accuracy and the percentage of trials where participants reported a strong, moderate, weak, or wrong performance is calculated.

4.2 Results

4.2.1 Participant's tasks preferences and difficulty ratings

Figure 4.2. illustrates the participants' responses to the question of *"If you were to choose one of the tasks below to do on a daily basis to use a computer, which one would you pick?"*when given the three different sets of options. For each scenario, the number of participants who chose each option as their favorite is indicated.

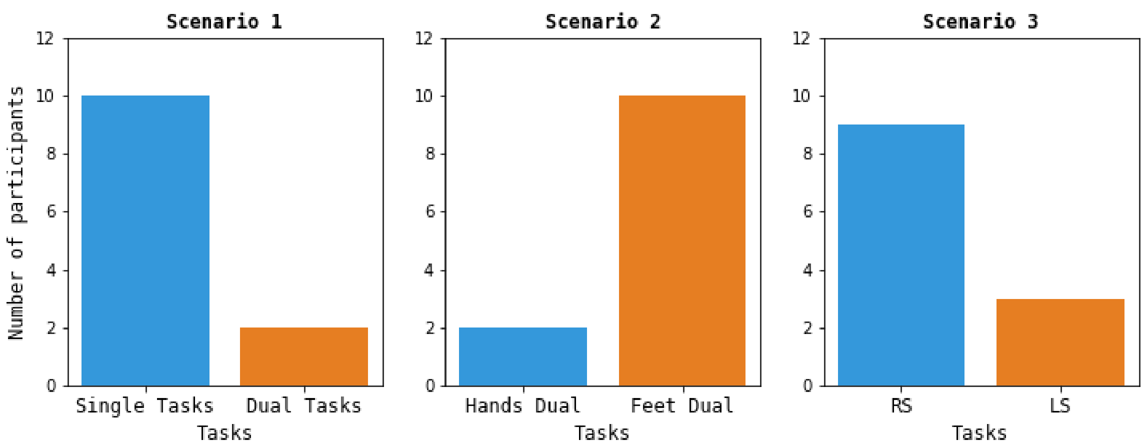

Figure 4.2: Participant responses to the task preference question.

Table 4.1: Task difficulty ratings

Task difficulty ratings				
	Recording after			Task Average
	Block 1	Block 4	Block 6	
REST	1.2 (0.4)	1.3 (0.5)	1.3 (0.5)	**1.2 (0.3)**
R	2.2 (1.1)	2.1 (0.9)	1.8 (1.1)	**2.0 (0.9)**
L	2.3 (0.8)	2.2 (0.9)	1.9 (1.0)	**2.1 (0.8)**
F	1.9 (0.9)	2.0 (1.1)	1.8 (1.1)	**1.9 (1.0)**
SI	2.4 (1.2)	2.0 (0.9)	1.9 (0.7)	**2.1 (0.9)**
$\mathbf{R_{SI}}$	3.3 (1.4)	2.8 (1.0)	2.4 (1.0)	**2.8 (1.0)**
$\mathbf{L_{SI}}$	3.4 (1.1)	2.8 (0.9)	2.5 (0.9)	**2.9 (0.9)**
$\mathbf{F_{SI}}$	2.8 (1.3)	2.6 (1.2)	2.3 (1.1)	**2.6 (1.2)**
Block Average	**2.5 (1.0)**	**2.2 (0.9)**	**2.0 (0.9)**	

Also, participants rated the difficulty of each task on a five-point scale after completing Blocks 1, 4, and 6. The average difficulty ratings (across participants) and the calculated standard deviation (in parenthesis) are reported in Table 4.1. In this table, block average is the average perceived difficulty of all the tasks after that block, and task average is the average perceived difficulty of that specific task recorded at three different points during the experiment.

4.2.2 EEG classification

The results of the investigated 3- and 4-class scenarios are available in Tables 4.2. The precision,recall, and F1-score metrics reported in these tables are defined in equations 4.1, 4.2 and 4.3, respectively. Figure 4.3 shows the corresponding confusion matrices. All results are averaged across the twelve participants.

$$Precision = \frac{TP}{TP + FP} \tag{4.1}$$

$$Recall = \frac{TP}{TP + FN} \tag{4.2}$$

$$F1 - score = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \tag{4.3}$$

where

TP: Number of times a task was correctly labeled as being a member of a specific class,

FP: Number of times a task was incorrectly labeled as being a member of a specific class,

FN: Number of times a member of a specific class was incorrectly labeled as being a member of another class.

The results of the 7-class combinations where the tasks, one at a time, were removed from the dataset are reported in Table 4.3. The reported value is the average accuracy when that particular task was not included in the classification.

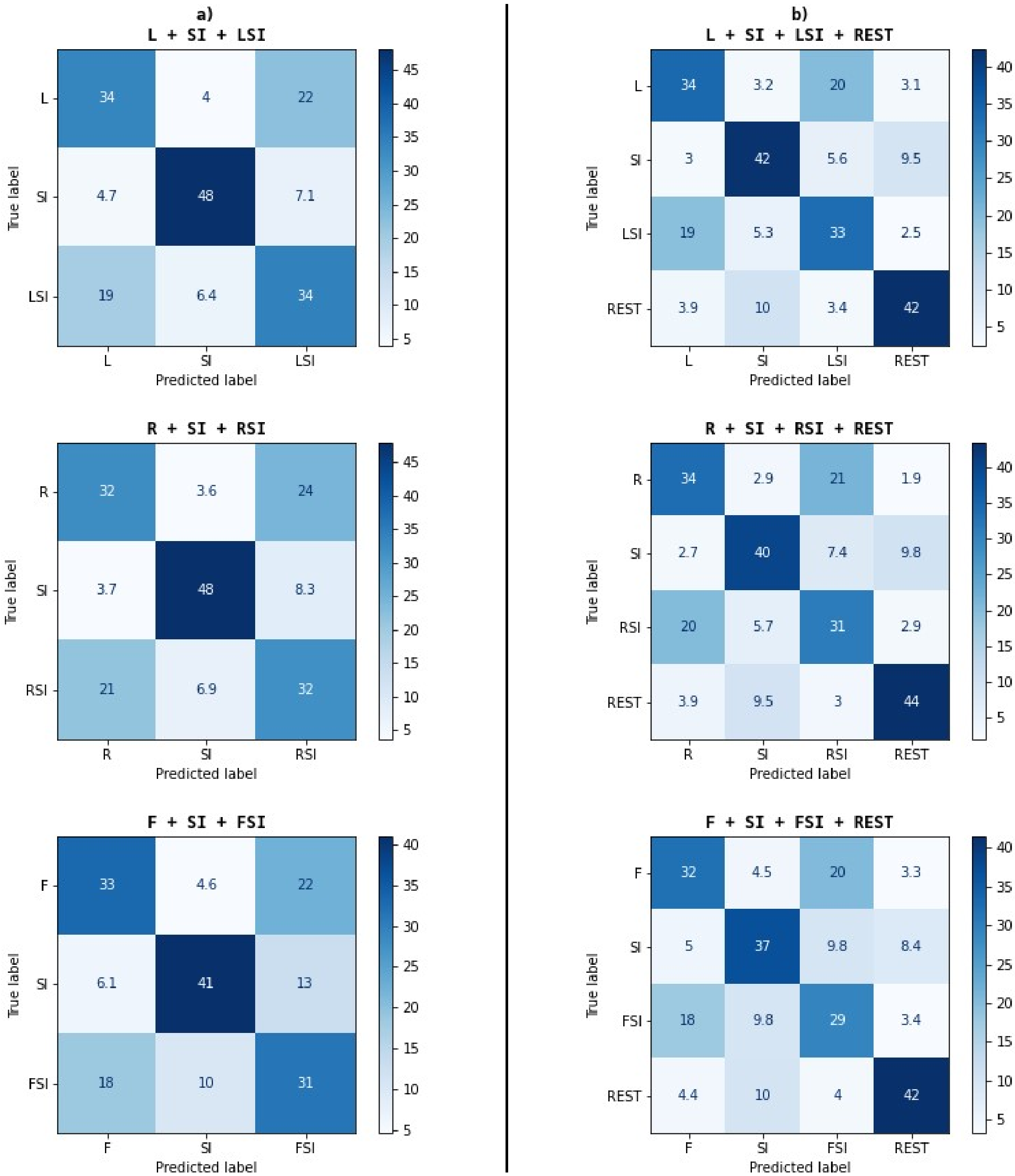

Figure 4.3: Confusion matrices for a) 3-class scenarios. b) 4-class scenarios.

Table 4.2: Performance evaluation of 3 and 4-class scenarios (%)

	3-Class scenarios (chance = 33.3 %)			4-Class scenarios (chance = 25.0%)		
	L + SI + L_{SI}			**L + SI + L_{SI} + REST**		
Overall Accuracy	64.1 (13.9)			63.0 (14.9)		
Task	Precision	Recall	F1-score	Precision	Recall	F1-score
L	58.5 (14.4)	56.1 (14.0)	57.1 (14.1)	56.2 (14.4)	56.4 (15.5)	56.2 (15.0)
SI	82.8 (19.1)	80.3 (17.8)	81.4 (18.3)	69.4 (17.4)	69.8 (17.3)	69.5 (17.2)
L_{SI}	53.6 (13.2)	57.0 (14.7)	55.1 (13.7)	53.6 (15.3)	55.0 (15.1)	54.2 (15.1)
REST	-	-	-	74.3 (16.4)	70.8 (15.5)	72.4 (15.8)
	R + SI + R_{SI}			**R + SI + R_{SI} + REST**		
Overall Accuracy	62.0 (11.3)			61.8 (12.6)		
Task	Precision	Recall	F1-score	Precision	Recall	F1-score
R	55.8 (15.4)	53.4 (16.4)	54.4 (15.8)	55.6 (14.2)	56.1 (15.4)	55.7 (14.6)
SI	82.1 (14.7)	80.0 (15.1)	80.9 (14.8)	68.8 (14.8)	66.7 (15.9)	67.5 (15.1)
R_{SI}	49.5 (10.2)	52.8 (10.2)	51.0 (10.0)	49.5 (10.6)	52.0 (11.3)	50.6 (10.9)
REST	-	-	-	75.5 (16.5)	72.5 (14.9)	73.9 (15.5)
	F + SI + F_{SI}			**F + SI + F_{SI} + REST**		
Overall Accuracy	59.3 (14.5)			58.1 (15.4)		
Task	Precision	Recall	F1-score	Precision	Recall	F1-score
F	57.5 (11.6)	55.1 (11.3)	56.1 (11.3)	54.2 (14.0)	53.3 (15.5)	53.5 (14.4)
SI	73.5 (18.5)	68.4 (18.3)	70.7 (18.3)	60.7 (21.1)	61.3 (21.0)	60.9 (20.9)
F_{SI}	47.5 (13.3)	52.4 (13.3)	49.7 (13.2)	46.7 (15.9)	48.7 (14.3)	47.6 (15.0)
REST	-	-	-	73.8 (17.1)	69.2 (16.4)	71.3 (16.6)

Table 4.3: Classification accuracies for the 7-class scenarios (%)

	7-Class scenarios (chance = 14.3%)							
	MI			**SI**	**DI**			**REST**
	L	**R**	**F**		**L_{SI}**	**R_{SI}**	**F_{SI}**	
Accuracy	53.8 (16.7)	52.8 (16.7)	52.2 (15.1)	51.6 (16.4)	54.2 (16.4)	55.4 (16.1)	55.1 (15.3)	49.9 (15.3)

Table 4.4: Performance evaluation metrics for the 8-class scenario (%)

8-Class Scenario								
Metric	**MI**			**SI**	**DI**			**REST**
	L	**R**	**F**		$\mathbf{L_{SI}}$	$\mathbf{R_{SI}}$	$\mathbf{F_{SI}}$	
Precision	49.3 (17.7)	52.3 (16.7)	48.2 (17.6)	52.2 (19.3)	45.1 (19.0)	43.4 (16.2)	41.4 (18.5)	69.0 (18.4)
Recall	50.3 (18.3)	51.9 (14.8)	50.0 (19.6)	49.3 (18.5)	46.2 (21.1)	44.0 (17.3)	41.3 (16.1)	64.8 (18.4)
F1-score	49.7 (17.9)	52.0 (15.6)	49.0 (18.5)	50.6 (18.8)	45.5 (19.9)	43.6 (16.6)	41.3 (17.3)	66.6 (18.1)

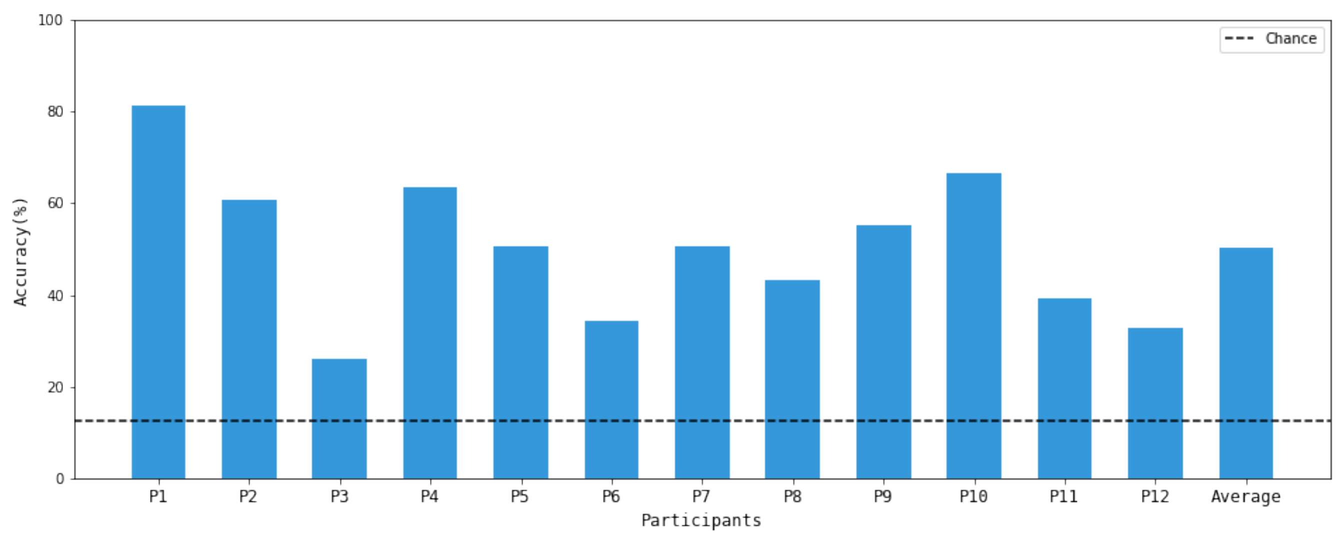

Figure 4.4: Individual participant accuracies for the 8-class scenario.

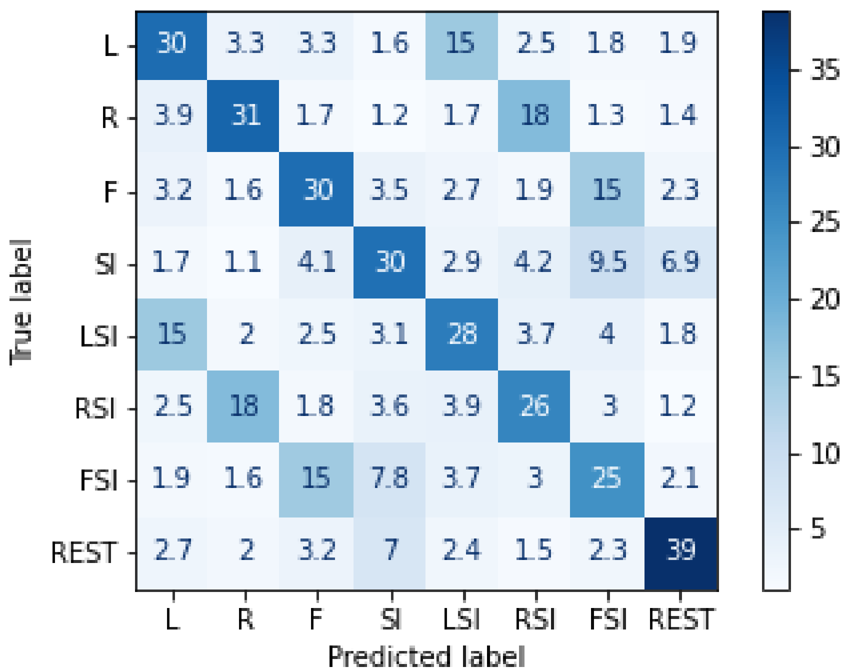

Figure 4.5: Confusion matrix for the 8-class scenario.

For the 8-class scenario, which includes all of the tasks considered in this study, an average accuracy of 50.5 ± 15.5% was achieved. Moreover, Table 4.4 provides the average performance evaluation metrics (precision, recall, and F1-score) for each of the eight tasks. The individual participant accuracies for the 8-class scenario are shown in Fig 4.4.

Figure 4.5 shows the average confusion matrix for the 8-class scenario and demonstrates the trends of common mistakes and errors for the classifier.

4.2.3 Correlation between classification accuracy and self-rating data

Fig 4.6 is a summary of the results of self-rating for each participant, reporting the percentage of trials for which they indicated their performance was “strong”, “moderate”, “weak”, or “wrong”.

None of the performance ratings demonstrated a significant correlation with the

8-class classification accuracy (p>0.23).

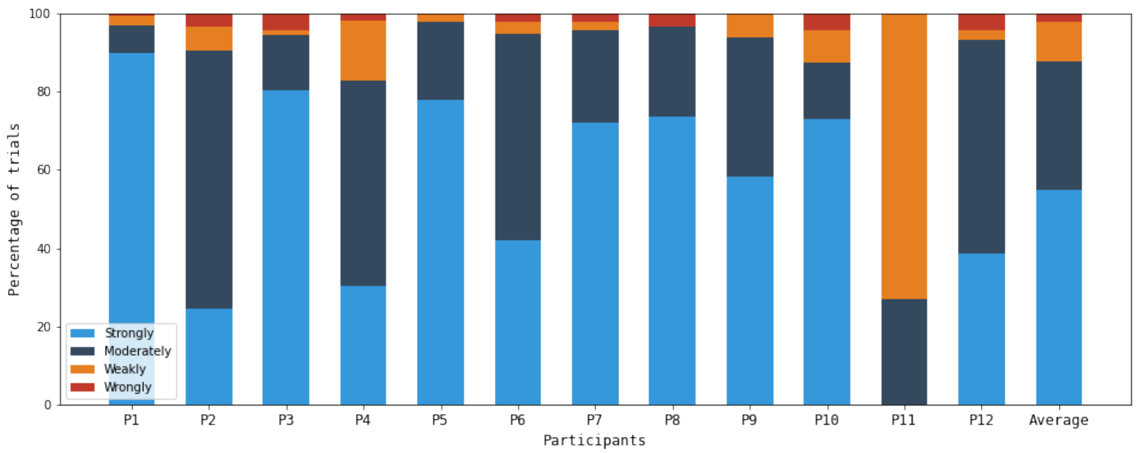

Figure 4.6: Results of self-rating for all of the participants.

4.3 Discussion

4.3.1 Task preferences and difficulty ratings

The importance of how participants perceive different tasks and their impact on their experience in using the active BCI has been discussed in previous chapters. In this experiment, we asked the participants to perform DI tasks which were expected to be more complex than single MI or SI tasks. Therefore, it is no surprise that 10 out of 12 participants found MI tasks preferable to DI tasks. Also, it was quite expected for participants to rate DI tasks as more difficult than MI and SI.

Encouragingly, however, a downward trend in participant difficulty ratings across the session was observed for all tasks (except REST), with the decrease in difficulty rating being larger for the DI tasks than for the single tasks. Specifically, while the average difficulty rating for the single tasks reduced from around 2.2 after block 1 to about 1.9 (=0.3) by the end of the session, the average DI task difficulty ratings changed from about 3.2 to 2.4 over the same period (=0.8). Therefore, it appears that the initial difficulty of performing DI tasks was considerably relieved as participants spent more time performing the tasks and gained more familiarity with them, and the difficulty ratings of DI ended up being relatively close to that of the single MI and SI tasks. Presumably, with additional practice/training, the perceived difficulty of performing the DI tasks could reduce even further.

All of the participants in this study were right-handed, and 9 out of 12 individuals preferred performing DI with their dominant hand (i.e., right) vs. their non-dominant hand. Moreover, the difficulty ratings appear to corroborate the responses to the task preference questions, which indicated that participants generally preferred the F_{SI} task to the R_{SI} and L_{SI} tasks - F_{SI} had a lower average difficulty rating (1.9/5) in comparison to the other two tasks (2.0/5 and 2.1/5, respectively). Overall, it seems that participants found F_{SI} to be a relatively more intuitive option.

4.3.2 EEG classification results

Dual vs. Single task classification

The first concern in using DI tasks for increasing the number of commands is the feasibility of differentiating the DI tasks and each of the constituent single tasks (e.g., L vs. SI vs. LSI). The results of the 3 and 4-class scenarios demonstrate that it is possible to design a classifier capable of differentiating an MI task, SI, and their corresponding dual task, with accuracies well above the chance level (i.e., 33.3 % and 25% for 3 and 4-class problems, respectively). This was true for all task combinations that were investigated. In all cases, the precision and recall metrics of the individual tasks were also all well above chance.

The results suggest that the DI and MI tasks are more often mistaken for each other than are the DI and SI tasks. This may indicate that the MI tends to generate stronger and clearer patterns of activity and could potentially overpower the SI component in the recorded neural signals of the DI tasks. Also, it is interesting to note that as the number of classes is increased, the classifier appears to improve at distinguishing the DI and single tasks (the single MI tasks are less often mislabelled as the corresponding DI task).

On the other hand, the SI component appears to be stronger in the generated brain activity for the F_{SI} task as compared to R_{SI} and L_{SI} since F_{SI} is more often incorrectly predicted as SI than are the other two DI tasks. As a result, the 3 and 4-class combinations that were based on F, SI, and F_{SI} appear to be slightly outperformed by the other two combinations in terms of accuracy.

It should also be noted that informal feedback received from the participants indicated that some individuals found it difficult to "turn on" the imagination of singing the song for the DI tasks and then "turn it off" for the single tasks. This could have negatively impacted the classification results by increasing the similarities of different tasks and their associated brain patterns. This could be something that, with training and practice, could become easier for participants.

The value of incorporating REST, or a "no control" state, into an active BCI has been discussed in previous chapters. In short, having the BCI be able to recognize an idle state would mean that the user would not be required to perform a specific task during every control period (they could, for example, just do nothing if they do not wish to send a command). The 4-, 7-, and 8-class analyses suggest that all of the single tasks and their dual versions are distinguishable from REST with accuracies well above chance, with the MI and DI tasks appearing to be more differentiable from REST than is SI (i.e., the SI task is more often mislabelled as REST than are MI and DI).

Increasing the number of commands

In Study 1, the possibilities of designing a 6-class active BCI that incorporated the SI into MI-BCI were investigated. Here, we explored the possibility of increasing the number of commands beyond that number. Specifically, 7 and 8-class scenarios were explored to evaluate the potential utility of DI tasks in multi-class active BCI. For the 8-class scenario, a very encouraging average accuracy of 50.5%, which is more than four times the level of chance, was achieved.

In addition, all possible 7-class combinations (obtained by excluding one of the eight tasks in each case) were considered. The highest 7-class accuracy yielded a 4.9% improvement as compared to the 8-class scenario, at the expense of one less command. This accuracy was obtained by excluding the R_{SI} task, which was not surprising given that the confusion matrix for the 8-class scenario indicates that the most common error was misclassifying R as R_{SI} or vice versa. Also, R_{SI} and F_{SI} were most frequently misclassified. Consequently, 7-class scenarios in which one of these two tasks was removed led to higher classification accuracies.

While the results obtained here are very promising and show the potential of using DI tasks in active BCI, the accuracies obtained for all of the 7- and 8-class combinations explored were below the 70% threshold, which is often said to be required for efficient BCI communication [34]. Nevertheless, it may be reasonable to assume that as participants get more familiar and comfortable with the DI tasks, the classification accuracy will be enhanced. Also, an improvement in feature extraction, selection, or classification algorithm could further enhance the performance of these paradigms.

4.3.3 Self-rating data

The observed lack of correlation between the participant's performance ratings and their 8-class classification accuracies further supports the idea that subjective data is not a reliable source of information for estimating the classifier performance.

4.4 Conclusion

The results of Study 2 indicate that DI tasks could be an effective approach for increasing the number of distinguishable commands in an active BCI to seven or eight. Specifically, an average accuracy of around 50% was achieved for the 8-class scenario, which is more than four times the corresponding chance level of 12.5%. An increase in the number of BCI commands would lead to increasing the information transfer rate of BCI. This can significantly augment the practicality and functionality of the system and, as a result, the quality of life for users.

Chapter 5

Conclusions

5.1 Contributions

The results of this study demonstrated an encouraging potential for using SI and DI tasks for enhancing MI-BCI design. Specifically, the following major points were determined from each of the studies:

- In Study 1, considering the following set of mental tasks (SI, L, R, F, T, REST):

    1. Maximum accuracies of 90.5%, 80.2%, 73.7%, and 67.1% were achieved for the 2, 3, 4, and 5-class combinations that incorporated SI, respectively. These values were comparable to those achieved for the combinations without SI.
    2. Maximum accuracies of 60.7% were achieved for 6-class combinations of tasks that incorporated SI, supporting the possibility of increasing the number commands for active BCIs.

- In Study 2, considering the following set of mental tasks (L, R, F, SI, L_{SI}, R_{SI}, F_{SI}, REST), where L_{SI}, R_{SI}, and F_{SI} are dual imagery tasks:

    1. 3-class combinations consisting of dual imagery tasks and their constituent single tasks were classified with accuracies higher than chance (i.e., 64.1%, 62.0%, 59.3% for the R, L, and F tasks, respectively).
    2. 4-class combinations consisting of dual imagery tasks, their constituent single tasks, and REST were classified with accuracies higher than chance (i.e., 63.0%, 61.8%, 58.1% for the R, L, and F tasks, respectively).
    3. 7- and 8-class scenarios consisting of various dual and single imagery task combinations were classified with accuracies as high as 55.4% and 50.5%, respectively. These results support the possibility of increasing the number of commands for active BCIs.

5.2 Limitations and future work

While the results of this study are encouraging, there are some limitations worth noting when interpreting the outcomes:

1. Since this study only involved healthy participants, unintended muscle activity artifacts could have impacted the classifier performance. Also, it is not known if the brain activity patterns generated by healthy participants are indicative of the patterns that would be generated by the target population of users. Hence, future work may address these limitations by recruiting participants from the target population of users and/or using EMG sensors to monitor unintended muscle activities.

2. Due to the limited sample size and the offline nature of the study, it is unclear how these results would translate to online scenarios. Future work may focus on utilizing the discussed tasks and scenarios in a real-time setting.

3. To better understand and evaluate the differences between the two types of SI tasks considered in Study 1, two groups of participants should be recruited, where each group would only perform one kind of SI. This would help avoid the possibility of "crosstalk" between the two SI tasks and alleviate the difficulty participants reported in differentiating the two sets of instructions.

4. This study was conducted under controlled research conditions. Therefore, it is not clear how the results would translate to everyday, real-world scenarios in which BCIs would actually be used. An online study involving the discussed mental tasks and scenarios could provide further information regarding their efficiency and practicality.

Furthermore, building on what was covered in this thesis, future work should focus on examining the possibilities of improvement through more optimized and even customized classification algorithms. These algorithms should be implemented in an online BCI to verify that the offline results translate.

Moreover, a further assessment of participants' perception of the introduced tasks should be carried out in multiple sessions to better evaluate the long-term suitability of the proposed novel paradigms. An investigation of how user training affects both classification accuracy and perception of the tasks should be conducted.

Crucially, the BCI design process must include individuals from the target population of users.

Bibliography

[1] R A Adcock et al. “Functional neuroanatomy of executive processes involved in dual-task performance”. en. In: *Proc. Natl. Acad. Sci. U. S. A.* 97.7 (2000), pp. 3567–3572.

[2] Kai Keng Ang et al. “Filter Bank Common Spatial Pattern algorithm on BCI Competition IV datasets 2a and 2b”. en. In: *Front. Neurosci.* 6 (2012), p. 39.

[3] Hubert Banville, Rishabh Gupta, and Tiago H Falk. “Mental task evaluation for hybrid NIRS-EEG brain-computer interfaces”. In: *Comput. Intell. Neurosci.* 2017 (2017), pp. 1–24.

[4] R Bellman and R Kalaba. “On adaptive control processes”. In: *IRE Trans. Autom. Control* 4.2 (1959), pp. 1–9.

[5] Paola Bucci and Silvana Galderisi. “Physiologic basis of the EEG signal”. In: *Standard Electroencephalography in Clinical Psychiatry.* Chichester, UK: John Wiley & Sons, Ltd, 2011, pp. 7–12.

[6] Soren Moller Christensen, Nicklas Stubkjar Holm, and Sadasivan Puthusserypady. “An improved five class MI based BCI scheme for drone control using filter bank CSP”. In: *2019 7th International Winter Conference on Brain-Computer Interface (BCI).* IEEE, 2019.

[7] Eleanor Curran et al. “Cognitive tasks for driving a brain-computer interfacing system: a pilot study”. en. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 12.1 (2004), pp. 48–54.

[8] Chris Ding and Hanchuan Peng. “Minimum redundancy feature selection from microarray gene expression data”. en. In: *J. Bioinform. Comput. Biol.* 03.02 (2005), pp. 185–205.

[9] Monica-Claudia Dobrea and Dan Marius Dobrea. "The selection of proper discriminative cognitive tasks — A necessary prerequisite in high-quality BCI applications". In: *2009 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies*. IEEE, 2009.

[10] Suman Dutta, Mandeep Singh, and Amod Kumar. "Classification of non-motor cognitive task in EEG based brain-computer interface using phase space features in multivariate empirical mode decomposition domain". In: *Biomed. Signal Process. Control* 39 (2018), pp. 378–389.

[11] Mohammed Zeki Al Faiz and Ammar A Al-Hamadani. "Online brain computer interface based five classes EEG to control humanoid robotic hand". In: *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, 2019.

[12] Farhad Faradji, Rabab K Ward, and Gary E Birch. "A brain-computer interface based on mental tasks with a zero false activation rate". In: *2009 4th International IEEE/EMBS Conference on Neural Engineering*. IEEE, 2009.

[13] B Farnsworth. *The Complete Pocket Guide; IMotions, Global HQ*. Copenhagen, Denmark, 2019.

[14] Jerome H Friedman. "Regularized Discriminant Analysis". In: *J. Am. Stat. Assoc.* 84.405 (1989), p. 165.

[15] Elisabeth V C Friedrich, Christa Neuper, and Reinhold Scherer. "Whatever works: A systematic user-centered training protocol to optimize brain-computer interfacing individually". en. In: *PLoS One* 8.9 (2013), e76214.

[16] Elisabeth V C Friedrich, Reinhold Scherer, and Christa Neuper. "Long-term evaluation of a 4-class imagery-based brain-computer interface". en. In: *Clin. Neurophysiol.* 124.5 (2013), pp. 916–927.

[17] Elisabeth V C Friedrich, Reinhold Scherer, and Christa Neuper. "Stability of event-related (de-) synchronization during brain-computer interface-relevant mental tasks". en. In: *Clin. Neurophysiol.* 124.1 (2013), pp. 61–69.

[18] Elisabeth V C Friedrich, Reinhold Scherer, and Christa Neuper. "The effect of distinct mental strategies on classification performance for brain-computer interfaces". en. In: *Int. J. Psychophysiol.* 84.1 (2012), pp. 86–94.

[19] Elisabeth V C Friedrich et al. "Impact of auditory distraction on user performance in a brain-computer interface driven by different mental tasks". en. In: *Clin. Neurophysiol.* 122.10 (2011), pp. 2003–2009.

[20] K Fukunaga. *Introduction to Statistical pattern recognition.* Academic Press Inc, 1990.

[21] Tao Geng et al. "A novel design of 4-class BCI using two binary classifiers and parallel mental tasks". en. In: *Comput. Intell. Neurosci.* 2008 (2008), p. 437306.

[22] Moritz Grosse, - Wentrup, and Martin Buss. "Multiclass common spatial patterns and information theoretic feature extraction". In: *IEEE Transactions on Biomedical Engineering* 55.8 (1991).

[23] Huma Hamid et al. "Analyzing classification performance of fNIRS-BCI for gait rehabilitation using deep neural networks". en. In: *Sensors (Basel)* 22.5 (2022), p. 1932.

[24] Sébastien Hétu et al. "The neural network of motor imagery: an ALE meta-analysis". en. In: *Neurosci. Biobehav. Rev.* 37.5 (2013), pp. 930–949.

[25] L M Hirshfield et al. "This is your brain on interfaces: enhancing usability testing with functional near- infrared spectroscopy". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 2011, pp. 373–382.

[26] Tin Ho. "Random Decision Forests (PDF)". In: *Proceedings of the 3rd International Conference on Document Analysis and Recognition.* Montreal, QC, 1995, pp. 278–282.

[27] Richard W Homan, John Herman, and Phillip Purdy. "Cerebral location of international 10–20 system electrode placement". en. In: *Electroencephalogr. Clin. Neurophysiol.* 66.4 (1987), pp. 376–382.

[28] Susanne M Jaeggi et al. "Does excessive memory load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: functional magnetic resonance imaging study". en. In: *Neuroimage* 19.2 (2003), pp. 210–225.

[29] G James. *An introduction to statistical learning.* California: Springer, 2013.

[30] Jun Jiang et al. "A novel Morse code-inspired method for multiclass motor imagery brain-computer interface (BCI) design". en. In: *Comput. Biol. Med.* 66 (2015), pp. 11–19.

[31] M J Khan. “Drowsiness detection in dorsolateral-prefrontal cortex using fNIRS for a passive-BCI”. In: *International Conference on Control, Automation and Systems (ICCAS)* 15 (2015).

[32] M Jawad Khan et al. “Drowsiness detection using fNIRS in different time windows for a passive BCI”. In: *2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE, 2016.

[33] Keun-Tae Kim, Tom Carlson, and Seong-Whan Lee. “Design of a robotic wheelchair with a motor imagery based brain-computer interface”. In: *2013 International Winter Workshop on Brain-Computer Interface (BCI)*. IEEE, 2013.

[34] A Kübler et al. “Brain-computer communication: self-regulation of slow cortical potentials for verbal communication”. en. In: *Arch. Phys. Med. Rehabil.* 82.11 (2001), pp. 1533–1539.

[35] Sarah Laszlo et al. “A direct comparison of active and passive amplification electrodes in the same amplifier system”. en. In: *J. Neurosci. Methods* 235 (2014), pp. 298–307.

[36] Liu, X., Makeyev, O., & Besio, W. “Improved Spatial Resolution of Electroencephalogram Using Tripolar Concentric Ring Electrode Sensors”. en. In: *Journal of Sensors* (2020).

[37] Ritik Looned et al. “Assisting drinking with an affordable BCI-controlled wearable robot and electrical stimulation: a preliminary investigation”. en. In: *J. Neuroeng. Rehabil.* 11.1 (2014), p. 51.

[38] F Lotte. “A review of classification algorithms for EEG-based brain computer interfaces: a 10 year update”. In: *Journal of Neural Engineering* 15 ().

[39] Na Lu et al. “A deep learning scheme for motor imagery classification based on restricted Boltzmann machines”. en. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 25.6 (2017), pp. 566–576.

[40] Francine Malouin et al. “The Kinesthetic and Visual Imagery Questionnaire (KVIQ) for assessing motor imagery in persons with physical disabilities: a reliability and construct validity study”. en. In: *J. Neurol. Phys. Ther.* 31.1 (2007), pp. 20–29.

[41] D J Mcfarland et al. *Mu and beta rhythm topographies during motor imagery and actual movements Brain Topography*. 2000.

[42] Dennis J McFarland et al. "Emulation of computer mouse control with a non-invasive brain-computer interface". en. In: *J. Neural Eng.* 5.2 (2008), pp. 101–110.

[43] Hadi Mohammadpour and Sarah D Power. "Investigating the addition of singing imagery as a control task in motor imagery BCI". In: *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2022.

[44] M K I Molla et al. *Trial regeneration with subband signals for motor imagery classification in BCI paradigm IEEE Access*. 2021.

[45] Th Mulder. "Motor imagery and action observation: cognitive tools for rehabilitation". en. In: *J. Neural Transm. (Vienna)* 114.10 (2007), pp. 1265–1278.

[46] Andrew Myrden and Tom Chau. "Effects of user mental state on EEG-BCI performance". en. In: *Front. Hum. Neurosci.* 9 (2015), p. 308.

[47] B Obermaier et al. "Information transfer rate in a five-classes brain-computer interface". en. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 9.3 (2001), pp. 283–288.

[48] Andrea Orlandi, Elisa Arno, and Alice Mado Proverbio. "The effect of expertise on kinesthetic motor imagery of complex actions". en. In: *Brain Topogr.* 33.2 (2020), pp. 238–254.

[49] Jerrin Thomas Panachakel and Angarai Ganesan Ramakrishnan. "Decoding covert speech from EEG-A comprehensive review". en. In: *Front. Neurosci.* 15 (2021), p. 642251.

[50] Hanchuan Peng, Fuhui Long, and Chris Ding. "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy". en. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 27.8 (2005), pp. 1226–1238.

[51] K Permana, S K Wijaya, and P Prajitno. "Controlled wheelchair based on brain computer interface using Neurosky Mindwave Mobile 2". In: *PROCEEDINGS OF THE 4TH INTERNATIONAL SYMPOSIUM ON CURRENT PROGRESS IN MATHEMATICS AND SCIENCES (ISCPMS2018)*. AIP Publishing, 2019.

[52] G Pfurtscheller and Da Silva. "Event-related EEG/MEG synchronization and desynchronization: basic principles Clin". In: *Clin. Neurophysiol* (1999).

[53] M Rashid et al. "The classification of motor imagery response: an accuracy enhancement through the ensemble of random subspace k-nn PeerJ Comput". In: *PeerJ Comput. Sci* (2021).

[54] G Reshmi and A Amal. "Design of a BCI system for piloting a wheelchair using five class MI based EEG". In: *2013 Third International Conference on Advances in Computing and Communications*. IEEE, 2013.

[55] Reinhold Scherer et al. "Brain–computer interfacing: more than the sum of its parts". en. In: *Soft Comput.* 17.2 (2013), pp. 317–331.

[56] Robin Tibor Schirrmeister et al. "Deep learning with convolutional neural networks for EEG decoding and visualization". en. In: *Hum. Brain Mapp.* 38.11 (2017), pp. 5391–5420.

[57] Francisco Sepulveda et al. "A comparison of mental task combinations for asynchronous EEG-based BCIs". en. In: *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2007 (2007), pp. 5055–5058.

[58] Amardeep Singh et al. "A comprehensive review on critical issues and possible solutions of motor imagery based electroencephalography brain-computer interface". en. In: *Sensors (Basel)* 21.6 (2021), p. 2173.

[59] David Steyrl et al. "Random forests in non-invasive sensorimotor rhythm brain-computer interfaces: a practical and convenient non-linear classifier". en. In: *Biomed. Tech. (Berl.)* 61.1 (2016), pp. 77–86.

[60] Irene Sturm et al. "Interpretable deep neural networks for single-trial EEG classification". en. In: *J. Neurosci. Methods* 274 (2016), pp. 141–145.

[61] Yousef Rezaei Tabar and Ugur Halici. "A novel deep learning approach for classification of EEG motor imagery signals". en. In: *J. Neural Eng.* 14.1 (2017), p. 016003.

[62] Ryo Takahashi and Hisaya Tanaka. "Performance improvement of motor-imagery BCI using multi-mental tasks". en. In: *International Symposium on Affective Science and Engineering* ISASE2019.0 (2019), pp. 1–4.

[63] H Al-Turabi and H Al-Junaid. "Brain computer interface for wheelchair control in smart environment". In: *Smart Cities Symposium 2018*. Institution of Engineering and Technology, 2018.

[64] Beatriz Vasconcelos et al. “The arch electrode: A novel dry electrode concept for improved wearing comfort”. en. In: *Front. Neurosci.* 15 (2021), p. 748100.

[65] Aleksandra Vuckovic and Bethel A Osuagwu. “Using a motor imagery questionnaire to estimate the performance of a Brain-Computer Interface based on object oriented motor imagery”. en. In: *Clin. Neurophysiol.* 124.8 (2013), pp. 1586–1595.

[66] Cong Wang et al. “Motor imagery BCI-based robot arm system”. In: *2011 Seventh International Conference on Natural Computation*. IEEE, 2011.

[67] Jonathan R Wolpaw and Elizabeth Winter Wolpaw. “Brain–computer interfaces: Something new under the sun”. In: *Brain–Computer InterfacesPrinciples and Practice*. Oxford University Press, 2012, pp. 3–12.

[68] Kunyu Xu, Yu-Yu Huang, and Jeng-Ren Duann. “The sensitivity of single-trial mu-suppression detection for motor imagery performance as compared to motor execution and motor observation performance”. en. In: *Front. Hum. Neurosci.* 13 (2019), p. 302.

[69] Yang Yu et al. “An asynchronous control paradigm based on sequential motor imagery and its application in wheelchair navigation”. en. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 26.12 (2018), pp. 2367–2375.

[70] J Zoltan, Michael S Koles, and Steven Z Lazar. “Spatial patterns underlying population differences in the background EEG”. In: *Brain Topography* 2.4 (1990), pp. 275–284.