A Neurofeedback-Based Near-Infrared Spectroscopy Brain-Computer Interface

by

Sabine Angelika Weyand M.A.Sc

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Graduate Department of Biomaterials and Biomedical Engineering
University of Toronto

A Neurofeedback-Based Near-Infrared Spectroscopy Brain-Computer Interface

Sabine Angelika Weyand

Doctor of Philosophy

Graduate Department of Biomaterials and Biomedical Engineering
University of Toronto

2015

Abstract

Near-infrared spectroscopy (NIRS) brain-computer interfaces (BCIs) enable individuals to interact with their environment using only cognitive activities. This thesis investigates the development of a more user-friendly, intuitive, and easy to use NIRS-BCI through six research objectives: exploring prescribed and personalized mental task frameworks offline, using researcher-selected tasks to move beyond the binary paradigm, exploring correlations of user characteristics with accuracy, comparing user-selected personalized tasks to prescribed tasks online, weaning off mental tasks to achieve voluntary self-regulation, and applying personalized frameworks to a client case study.

Firstly, personalized tasks outperformed prescribed tasks in a five-session offline study conducted on ten able-bodied participants. Specifically, user-selected tasks resulted in significantly higher ease-of-use, while researcher-selected tasks resulted in significantly higher accuracies. The same data were used to show that researcher-selected personalized mental tasks enabled classification in some users beyond a binary BCI paradigm. Accuracy was strongly positively correlated with perceived ease of session, ease of concentration, and enjoyment, but strongly negatively correlated with verbal IQ. In a second study, when comparing two able-bodied groups online (N = 9 and N = 10), the usability of user-selected personalized

mental tasks exceeded prescribed mental tasks without a decrease in accuracy. Expanding on this study, the nine able-bodied subjects who used user-selected tasks took part in an additional ten sessions and were weaned off mental tasks to achieve online voluntary self-regulatory control of a BCI using a neurofeedback-based paradigm. Participants indicated that they found self-regulation to be more intuitive and easier to use than mental tasks. Finally, user- and researcher-selected frameworks were applied to a client with undiagnosed motor impairments, unveiling a host of neuropsychological challenges to BCI control. Overall, this thesis advances the field of knowledge of NIRS-BCIs, specifically with respect to usability.

Acknowledgments

I would like to express my deepest appreciation to my supervisors Dr. Tom Chau and Dr. Kaori Takehara-Nishiuchi for their help, guidance, support and wisdom over the past few years. I would also like to thank my committee members, Dr. Milos Popovic and Dr. Jeff Daskalakis, whose guidance and encouragement have been greatly appreciated. I would like to acknowledge my external examiners Dr. Patricia Trbovich and Dr. Tanja Schultz for their valuable feedback. I would also like to express my gratitude to all the members of the PRISM lab for their encouragement, comradery, and help.

I would like to acknowledge the financial support that I have received over the years from the Holland Bloorview Kids Rehabilitation Hospital and the Natural Sciences and Research Council of Canada.

Finally, I would like to thank my family. To my husband, Michael, thank you for your patience and support. I will forever be indebted to you for reading the thousands of words I have put on paper over the last few years and for providing me with words of encouragement when I needed it most. To my parents, Angelika and Jurgen, I would like to thank you for all the opportunities you have given me. Your strength, perseverance, support, and unconditional love inspire me. I know that I can always count on you and for that I am truly grateful.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BCI	Brain-computer interface
BOLD	Blood oxygen level dependent
CI	Confidence interval
DPF	Differential pathlength factor
ECoG	Electrocorticography
EEG	Electroencephalography
FCBF	Fast correlation based filter
Hb	Deoxygenated haemoglobin
HbO	Oxygenated haemoglobin
HT	Happy thoughts
IEEE	Institute of Electrical and Electronics Engineers
IQ	Intelligence quotient
LDA	Linear discriminant analysis
MM	Mental math
MRI	Magnetic resonance imaging
NASA-TLX	National Aeronautics and Space Administration Task Load Index
NIRS	Near-infrared spectroscopy
NSERC	Natural Sciences and Engineering Research Council
OVO	One-vs-one
PET	Positron emission tomography
PFC	Prefrontal cortex

PRISM	Paediatric Rehabilitation Intelligent Systems Multidisciplinary
PWAR	Pair-wise accuracy rankings
PWAR-RS	Researcher-selected pair-wise accuracy rankings
PWAR-US	User-selected pair-wise accuracy rankings
RAM	Random-access memory
RF	Relaxing with focus
RR	Unconstrained rest
RS	Relaxing with slow counting
SFFS	Sequential forward floating search
tHb	Total haemoglobin
VP	Visual prompt
WAIS	Wechsler intelligent scale for adults
WG	Word generation
WS-scores	Weighted slope scores
WS-US	User-selected weighted slope score

Chapter 1: Introduction

1.1 Brain-Computer Interfaces

Brain-computer interfaces (BCIs) allow individuals to interact with their environment using only cognitive activities (Elisabeth V C Friedrich, Scherer, and Neuper 2012; Ang, Yu, and Guan 2012; S. M. Coyle, Ward, and Markham 2007). BCIs can serve as a conduit to communication or mobility for individuals with severe motor impairments resulting from amyotrophic lateral sclerosis, spinal cord injuries, brain stem stroke, muscular dystrophy or other debilitating conditions (Elisabeth V C Friedrich, Scherer, and Neuper 2012; Ayaz et al. 2007; Niels Birbaumer 2006; J. Wolpaw et al. 2000; Sitaram et al. 2007). BCIs can also be used by able-bodied individuals for productivity, gaming, entertainment, brain training, meditation and to accelerate learning (Elisabeth V C Friedrich, Scherer, and Neuper 2012; J. Wolpaw et al. 2000).

The basic components of a BCI are: the physiological input, the signal processing unit, the classifier, and the output. The input to the BCI can be further categorized into the access modality, which refers to how the physiological signal is collected, and the access pathway, which refers to how a change in the signal is evoked (Blain, Mihailidis, and Chau 2008; K. Tai, Blain, and Chau 2008; Ayaz et al. 2009). The main focus of this thesis is on improving the BCI access pathway.

1.2 Access Modalities

Access modalities can be classified as invasive and non-invasive. The most common invasive BCIs use electrocorticography (ECoG) or intracortical recordings. In ECoG, the brain activity is measured using electrodes implanted on the brain surface, while in intracortical recordings, the electrodes are implanted inside the cortex. Invasive BCIs have the advantages of high signal-to-noise ratio and high spatial resolution. However, invasive BCIs have the major limitation of requiring surgical implantation, which results in a high risk of neural tissue damage and infection (Blain, Mihailidis, and Chau 2008; K. Tai, Blain, and Chau 2008; Ayaz et al. 2009; Morshed and Khan 2014).

The most common non-invasive BCIs used to date are electroencephalography (EEG), magnetic resonance imaging (MRI), and near-infrared spectroscopy (NIRS). EEG is the most researched of the non-invasive access modalities and involves measuring the electrical brain activity, recorded from the scalp, using surface electrodes. EEG has a high temporal resolution and low cost. However, EEG is prone to artefacts, has a low signal-to-noise ratio, and requires gel and cumbersome electrode fixation (K. Tai, Blain, and Chau 2008; Ayaz et al. 2009).

MRI measures the haemodynamic brain activity by ascertaining the concentration of deoxygenated haemoglobin (Hb) in the brain through the blood oxygen level dependent (BOLD) effect (Ang, Yu, and Guan 2012; S. M. Coyle, Ward, and Markham 2007). The neuronal haemodynamic response results from local dilation of arterioles and capillaries in regions of neural activation. The dilation causes an increase in cerebral blood flow that exceeds the metabolic demand and results in a regional increase in the concentration of oxygenated haemoglobin (HbO) and a decrease in the concentration of Hb, peaking approximately five seconds after activation (Ayaz et al. 2009; Toomim et al. 2005; Arno Villringer and Chance 1997). This phenomenon is known as neurovascular coupling (S. M. Coyle, Ward, and Markham 2007; Niels Birbaumer and Cohen 2007; Wolf et al. 2002), although other coupling trends have also been reported (Bauernfeind et al. 2008; A Villringer et al. 1993; Gert Pfurtscheller, Bauernfeind, et al. 2010; Quaresima et al. 2005; Y Hoshi et al. 1994; Koshino et al. 2011; Buckner, Andrews-Hanna, and Schacter 2008). MRI does not require electrode gel, and is not affected by electrical noise or blinking of the eyes. However, MRI machines are very large and expensive, requiring specialized building infrastructure (e.g., vibration, acoustic and electromagnetic shielding) and therefore, are not ideally suited as BCIs (Ayaz et al. 2009; Sitaram et al. 2007).

The access modality used in this research is NIRS (Ang, Yu, and Guan 2012; S. M. Coyle, Ward, and Markham 2007; Ayaz et al. 2007; S. Coyle et al. 2004). Similar to MRI, NIRS measures the haemodynamic activity of the brain. To make NIRS measurements, a near-infrared light source is placed on the surface of the skin; the light travels through the bone and the meninges to the cortex and is scattered back through the tissue in a banana shaped path, to a detector (S. M. Coyle, Ward, and Markham 2007). The modified Beer-Lambert’s law can be used to calculate the amount of Hb and HbO based on the amount of light absorbed (S. M. Coyle, Ward, and Markham 2007; J. Wolpaw et al. 2000; Niels Birbaumer and Cohen 2007; Delpy et al. 1988). Compared to EEG, NIRS does not require electrode gel and is not affected by electrical noise or blinking of the eyes (S. M. Coyle, Ward, and Markham 2007; Ayaz et al. 2009; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; T. Falk et al. 2011; M Izzetoglu et al. 2005). Additionally, compared to MRI, NIRS is relatively inexpensive and does not require large equipment (Ayaz et al. 2009; Sitaram et al. 2007). However, similar to MRI, there is an inherent haemodynamic delay. Unlike MRI, NIRS only provides information about cortical activation; the activity of deeper brain structures is not measurable by NIRS (S. M. Coyle, Ward, and Markham 2007; Ayaz et al. 2009; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; T. Falk et al. 2011; M Izzetoglu et al. 2005).

1.3 Access Pathways

BCIs can also differ in the access pathway, which refers to how the signal is evoked. Prior to this work, to the best of our knowledge, all active NIRS-BCI studies have used prescribed mental activation tasks to control the BCI, where the user performs specific tasks that result in predictable changes in haemodynamic activity (Nicolas-Alonso and Gomez-Gil 2012; Strait and Scheutz 2014; L. Schudlo, Weyand, and Chau 2014; Zephaniah and Kim 2014). The tasks used to control the BCI are chosen by researchers based on previous studies showing differentiability in the activation or deactivation caused by a specific set of tasks. By discriminating between the changes in the NIRS signal accompanying the performance of different tasks, control of the binary BCI is achieved. A number of different mental tasks have been used in past NIRS-BCI studies, including: mental math (Ang, Yu, and Guan 2012; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Ogata, Mukai, and Yagi 2007; Bauernfeind et al. 2008; Sarah D. Power, Kushki, and Chau 2012; Utsugi et al. 2007; L. C. Schudlo and Chau 2014), mental singing (Naito et al. 2007; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011), word generation (Herff, Heger, Putze, et al. 2013; Ogata, Mukai, and Yagi 2007; Utsugi et al. 2007), memory (Ayaz et al. 2007; Ogata, Mukai, and Yagi 2007; Utsugi et al. 2007), mental counting (Naseer and Hong 2013a), mental rotation (Herff, Heger, Putze, et al. 2013), concentration (K. Izzetoglu et al. 2011), motor imagery (S. M. Coyle, Ward, and Markham 2007; Sitaram et al. 2007; S. Coyle et al. 2004; Kanoh et al. 2009; Naseer and Hong 2013b), and rest (Ang, Yu, and Guan 2012; Ayaz et al. 2007; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012; Naseer and Hong 2013a; L. C. Schudlo and Chau 2014). The limitations of using prescribed mental tasks include the fact that they may not be the most appropriate task for each person, they do not take user preferences into account, and they can be cognitively demanding and unintuitive.

1.4 Signal Processing

Throughout this thesis, NIRS data were collected using a multi-channel frequency-domain NIRS system (Imagent Functional Brain Imaging System from ISS Inc., Champaign, IL (ISS Inc. 2012)). The NIRS system was used to measure the blood oxygen content from the prefrontal cortex (PFC) at nine discrete locations. A variety of signal processing methods were then applied. Briefly, the NIRS signal was passed through a digital low-pass filter to mitigate the effect of various sources of physiological noise. After filtering the data, the changes in concentrations of HbO, Hb, and total haemoglobin (tHb), were calculated using the modified Beer-Lambert's Law (S. M. Coyle, Ward, and Markham 2007; J. Wolpaw et al. 2000; Niels Birbaumer and Cohen 2007; Delpy et al. 1988). Next, temporal and spatial features were extracted from the data. The use of both temporal and spatial features was motivated by the potential for information gain (L. C. Schudlo, Power, and Chau 2013). Finally, feature selection was performed using either sequential forward floating search (SFFS) (L. C. Schudlo, Power, and Chau 2013; Pudil, Novovičová, and Kittler 1994; Jain and Zongker 1997; Kudo and Sklansky 2000) or a fast correlation based filter (FCBF) (Yu and Liu 2003; Koelstra et al. 2010; Chanel, Ansari-Asl, and Pun 2007).

1.5 Classification and Output

Since the goal of a BCI is to control the environment, classification of a user's brain signals must be performed. Classification of NIRS signals can either be performed offline, following the completion of data collection, or online, in real-time, as the data are being collected. In general, the aim of offline classification is to provide an estimate of how a classifier, trained on the data collected, would perform on similar future data. Offline classification also provides the ability to make adjustments to the analysis methods, such as extracting and selecting different features. In contrast, online classification involves training a classifier using previously collected data, and then predicting the class of new data as the task is being performed. Online classification can enable real-time control and provide the user with immediate performance feedback. We chose to employ Fisher's linear discriminant analysis (LDA) classifier since it has been used successfully in earlier NIRS-BCI studies (L. C. Schudlo, Power, and Chau 2013; L. C. Schudlo and Chau 2014; S. Power, Kushki, and Chau 2011; Sarah D Power, Kushki, and Chau 2012; Sarah D. Power, Kushki, and Chau 2012; Sarah Dianne Power and Chau 2013; Moghimi, Kushki, Power, et al. 2012; Faress and Chau 2013; Herff, Heger, Fortmann, et al. 2013; Herff, Heger, Putze, et al. 2013; Bauernfeind et al. 2011; Kelly Tai and Chau 2009) and it is computationally efficient. Moreover, based on preliminary data analysis, LDA was shown to outperform or perform on par with support vector machines, naïve Bayes, hidden Markov model, *k*-Nearest Neighbors, and decision trees.

1.6 Motivation

NIRS-BCIs still lack in several areas, including: user-friendliness, intuitiveness, low information transfer rates, large variability in performance between users, and a lack of research on target populations. Given these limitations, this research focuses on improving NIRS-BCI usability.

1.6.1 Personalized mental task frameworks

A personalized mental task paradigm is an alternative to the currently used prescribed mental task framework. It involves selecting a specific set of tasks for each user instead of the same tasks for every user. Personalized mental task frameworks can be further sub-categorized into user-centered and researcher-centered varieties. To the best of our knowledge, personalized tasks have not been studied with NIRS-BCIs; however, researcher-selected personalized tasks have been explored in MRI (Sorger et al. 2009) and EEG (Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012) BCI studies. Researcher-selected personalized mental tasks are motivated by a large inter-subject variability in performance. Therefore, choosing the best tasks for each individual can result in increased accuracies (Sorger et al. 2009; Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012). On the other hand, user-selected personalized tasks are motivated by a large inter-subject variability in task ease-of-use and the importance of user satisfaction in the adoption and use of assistive technology (Sorger et al. 2009; Elisabeth V C Friedrich, Scherer, and Neuper 2012; E. Curran et al. 2004; Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012). Therefore, allowing users to choose their own personalized mental tasks could potentially improve the ease-of-use and adoption of the BCI (Sorger et al. 2009; Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012). To facilitate a user-centered approach that allows one to strike a personal balance between usability and performance, two task measures were invoked, namely, a measure of usability and a measure of performance. It is noted that we are primarily focusing on the user-satisfaction dimension of usability (ISO 9241-11 1998). The usability of each task was based on each user’s subjective post-task ease-of-use ratings (Tedesco and Tullis 2006; Sauro and Dumas 2009). Two methods for displaying task performance were proposed in this thesis: pair-wise accuracy rankings (PWAR) and weighted slope scores (WS-scores). PWAR constitute a ranked list of the differentiability of each possible pair-wise combination of tasks. PWAR are motivated by providing users with detailed information as to which tasks are likely to result in high classification accuracies. WS-scores correspond to a ranked list of tasks that tend to consistently increase or decrease haemodynamic activity. WS-scores are motivated by being neurofeedback-centered, which allows for an intuitive and simplified choosing method that focuses on discernable differences in the neurofeedback between the selected tasks.

1.6.2 Moving beyond the binary paradigm

Multiclass BCIs (beyond binary) have the potential to provide users with more outputs, thereby increasing the rate of communication (Shin et al. 2013). However, as the number of classes increases, so will the difficulty in discriminating between each class. To date, limited research on multi-class NIRS-BCIs has been conducted. To the best of our knowledge, three studies have explicitly explored multi-class NIRS-BCIs over the prefrontal cortex (PFC) that could potentially be used to control a computer, namely, (Herff, Heger, Fortmann, et al. 2013), (Hirshfield et al. 2009), and (Sarah D Power, Kushki, and Chau 2012). None of the participants in these studies exceeded the 70% threshold, often cited as required for BCI control (Andrea Kübler, Neumann, et al. 2001). One potential method for improving the classification accuracies in multi-class NIRS-BCIs is the use of researcher-selected personalized mental tasks.

1.6.3 Correlation of NIRS-BCI accuracy with user characteristics

Although it is widely accepted that some BCI users perform better than others, the reason for this disparity is not well established. Specifically, the prediction of BCI accuracy based on user characteristics, such as demographic traits, IQ, and state of mind, are not well explored in NIRS-BCI literature. Determining the correlation between user characteristics and performance may help reduce some of the large inter-subject variability in classification accuracies.

1.6.4 Achieving BCI control using self-regulation

Currently, to the best of our knowledge, all NIRS-BCIs use mental tasks to elicit changes in regional haemodynamic activity. One of the limitations of using mental tasks is that they can be cognitively demanding and unintuitive. Voluntary self-regulation involves the acquisition of voluntary control over one's physiological signals without the need to perform a mental task. Self-regulation has the potential to result in a more intuitive, easier to use, and less mentally demanding access pathway. The field of voluntary self-regulation in BCIs is still in its infancy; however, to date, several researchers have shown the potential of voluntary self-regulation in EEG-BCIs with users gaining control of the 8-12 Hz mu rhythms (J. R. Wolpaw et al. 1997; J R Wolpaw, McFarland, and Vaughan 2000; Daly and Wolpaw 2008; E. A. Curran and Stokes 2003) and slow cortical potentials (SCPs) (Kotchoubey et al. 1996; Andrea Kübler, Neumann, et al. 2001; Daly and Wolpaw 2008; A Kübler et al. 1999; E. A. Curran and Stokes 2003; Niels Birbaumer 2006; N. Birbaumer et al. 1981). To the best of our knowledge, no studies have explored voluntary self-regulation for the purpose of controlling an NIRS-BCI. A personalized mental task framework may be a particularly appealing means of facilitating self-regulation, where users start with but are eventually weaned from individualized tasks.

1.6.5 Performance on target populations

Despite the fact that one of the greatest potential benefits of BCIs is the provision of control and communication for patients with motor impairments, the vast majority of current research is still being conducted on able-bodied subjects. It is acknowledged that research on able-bodied individuals is important; however, the conclusions may not always transfer to the patient population due to both predictable and unpredictable reasons. Therefore, studies on individuals with motor-impairments are critical to eventual clinical translation (Grosse-wentrup and Schölkopf 2013). To the best of our knowledge, only three NIRS-BCI studies have been conducted on individuals with motor impairments (Naito et al. 2007; Sarah Dianne Power and Chau 2013; Gallegos-Ayala et al. 2014). Despite these early studies on individuals with motor impairments, there is still a paucity of research exploring NIRS-BCI performance for individuals with various disabilities.

1.7 Objectives and Research Questions

The overall objective of this thesis is to advance the development of a more user-friendly, intuitive, practical, and easy to use NIRS-BCI.

The six secondary objectives (O) and the associated research questions (RQ) are:

O1. To compare four mental task methodological frameworks: a prescribed task framework, two user-selected personalized task frameworks, and a researcher-selected personalized task framework.

RQ1A. What is the effect of mental task framework on offline accuracies, ease-of-use, computational time, and length of training? Specifically, we consider a prescribed mental task framework, two personalized user-selected mental task frameworks, and a personalized researcher-selected mental task framework.

RQ1B. What is the effect of personalized user-selected mental task framework on user preference?

O2. To determine if it is possible to use researcher-selected personalized mental tasks to develop a more practical BCI and move beyond a binary BCI paradigm.

RQ2. What levels of offline classification accuracies can be attained for a 2-, 3-, 4-, and 5-class BCI using a personalized researcher-selected framework?

O3. To determine if various user characteristics correlate to BCI accuracy.

RQ3. Is there any correlation between classification accuracies and users' verbal IQ, self-reported tiredness, self-reported concentration, self-reported ease of performance, or self-reported enjoyment?

O4. To compare the usability and performance of a user-selected personalized mental task NIRS-BCI to a prescribed mental task NIRS-BCI using a two group online experimental design.

RQ4A. What is the effect of neurofeedback, performance, and ease-of-use informed choice of personalized mental tasks on the ease-of-use of a NIRS-BCI?

RQ4B. What is the effect of neurofeedback, performance, and ease-of-use informed choice of personalized mental tasks on the online accuracy of a NIRS-BCI?

O5. To determine the usability and performance of a NIRS-BCI that requires users to wean off mental tasks to achieve voluntary self-regulation.

RQ5A. What level of online classification accuracies can be attained for a binary switch NIRS-BCI when users are weaned off mental tasks and use a voluntary desire to modulate their haemodynamic activity (self-regulation)?

RQ5B. When comparing mental tasks and self-regulation, which method do users prefer in terms of mental work load, intuitiveness, and ease-of-use?

O6. To apply personalized mental task frameworks to the NIRS-BCI training of a client with motor impairments.

> **RQ6**. What level of accuracy can be attained when applying user- and researcher-selected personalized mental task frameworks to a client with undiagnosed motor impairments?

1.8 Thesis Organization

Chapters 2 through 6 of this thesis are reproduced verbatim from manuscripts. Chapter 2 addresses the first research objective and corresponding research questions, as described in section 1.7. Chapter 3 addresses both the second and third research objectives and corresponding research questions. Chapters 4, 5, and 6 address the fourth, fifth, and sixth research objectives and corresponding research questions, respectively. As each of these chapters is reproduced from stand-alone entities, certain information in the introduction and methods sections may be redundant. The final chapter of this thesis summarizes the major original contributions of this work. A graphical depiction of the relationship between the research objectives and chapters of this thesis is shown in Figure 1.

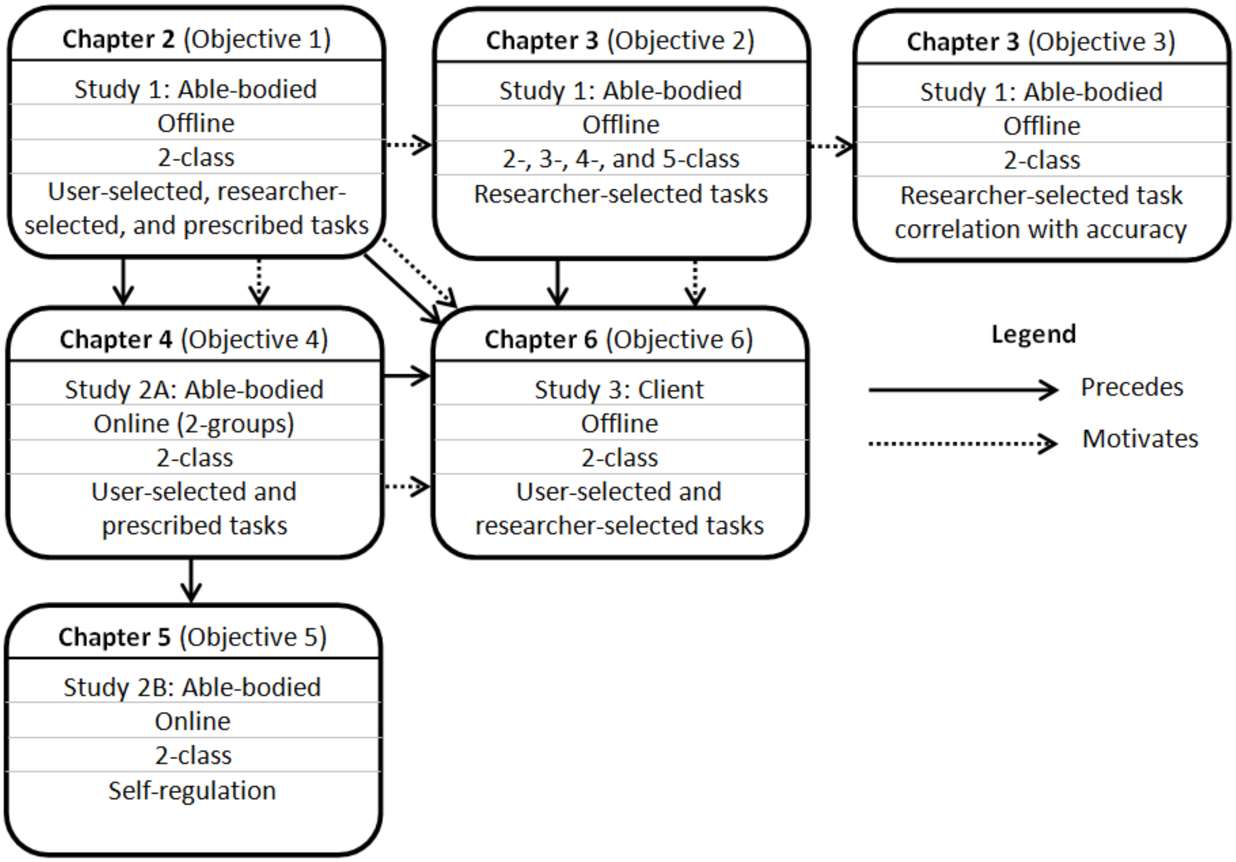

Figure 1. Graphical depiction of the relationship between the research objectives and chapters of this thesis.

Chapter 2: Offline Comparison of Four Methodological Frameworks

The entirety of this chapter is reproduced from the article “Exploring Methodological Frameworks For A Mental Task-Based Near-Infrared Spectroscopy Brain-Computer Interface”. This manuscript has been published in the Journal of Neuroscience Methods.

2.1 Abstract

Background: Near-infrared spectroscopy (NIRS) brain-computer interfaces (BCIs) enable users to interact with their environment using only cognitive activities. This paper presents the results of a comparison of four methodological frameworks used to select a pair of tasks to control a binary NIRS-BCI; specifically, three novel personalized task paradigms and the state-of-the-art prescribed task framework were explored.

New Methods: Three types of personalized task selection approaches were compared, including: user-selected mental tasks using weighted slope scores (WS-scores), user-selected mental tasks using pair-wise accuracy rankings (PWAR), and researcher-selected mental tasks using PWAR. These paradigms, along with the state-of-the-art prescribed mental task framework, where mental tasks are selected based on the most commonly used tasks in literature, were tested by ten able-bodied participants who took part in five NIRS-BCI sessions.

Results: The frameworks were compared in terms of their accuracy, perceived ease-of-use, computational time, user preference, and length of training. Most notably, researcher-selected personalized tasks resulted in significantly higher accuracies, while user-selected personalized tasks resulted in significantly higher perceived ease-of-use. It was also concluded that PWAR minimized the amount of data that needed to be collected; while, WS-scores maximized user satisfaction and minimized computational time.

Comparison with Existing Method: In comparison to the state-of-the-art prescribed mental tasks, our findings show that overall, personalized tasks appear to be superior to prescribed tasks with respect to accuracy and perceived ease-of-use.

Conclusions: The deployment of personalized rather than prescribed mental tasks ought to be considered and further investigated in future NIRS-BCI studies.

2.2 Introduction

2.2.1 Near-infrared spectroscopy brain computer interface

Brain-computer interfaces (BCIs) enable users to interact with their environment using only cognitive activities. A BCI consists of an input, signal-processing unit, classifier, and output. The input subsystem consists of the access modality, which refers to how the physiological signal is collected, and the access pathway, which refers to how a change in the signal is evoked (J. Wolpaw et al. 2000; Ayaz et al. 2009). For our study, near-infrared spectroscopy (NIRS) was chosen as the input access modality. NIRS is a safe optical neural imaging technique that can be used to measure haemodynamic brain activity (S. M. Coyle, Ward, and Markham 2007; Ayaz et al. 2007).

2.2.2 Mental task frameworks

To the best of our knowledge, all NIRS-BCI studies to date have used prescribed mental task access pathways, where researchers instruct all participants to perform a single set of predetermined tasks. As an alternative to prescribed mental tasks, personalized mental tasks have been proposed in magnetic resonance imaging (MRI) (Sorger et al. 2009) and electroencephalography (EEG) (Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012) BCI research. The use of personalized mental tasks is motivated by large inter-subject variability in task performance, task suitability, and task ease-of-use (Sorger et al. 2009; Elisabeth V C Friedrich, Scherer, and Neuper 2012; E. Curran et al. 2004; Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012).

Personalized mental task frameworks can be further sub-categorized into user-centered and researcher-centered. To the best of our knowledge, all MRI-BCI and EEG-BCI studies to date have focused on researcher-centered selection methods, which involve the researcher choosing each user's personalized mental tasks based on task performance. On the other hand, a user-centered approach, which involves each user choosing their own tasks, could also be explored. A user-centered design is motivated by the importance of user satisfaction in the adoption and use of assistive technology (Sorger et al. 2009; J. Wolpaw et al. 2000; Bos, Poel, and Nijholt 2011; Tan and Nijholt 2010).

2.2.3 Objectives

The goal of this research study was to evaluate four mental task frameworks: one prescribed framework, two user-centered frameworks, and a researcher-centered framework. Specifically, the frameworks were compared based on accuracy, perceived ease-of-use ratings, computational time, user preference, and required data for training.

2.3 Methods

It is noted that the data collected during this study were also analyzed to explore multi-class NIRS-BCI correlates. For more information on this work, please refer to (Weyand, Takehara-Nishiuchi, and Chau 2015a). As a result of observed participant head motion or loss of contact between the head and the detectors, up to 20 data points (a maximum of four per class) were discarded from participants 3 and 9.

2.3.1 Participants

Ten able-bodied subjects (4 male, 6 female) between the ages of 16 and 40 were recruited from the staff and students at Holland Bloorview Kids Rehabilitation Hospital (Toronto, Canada). Eight participants were right-handed according to the Edinburgh handedness test (Oldfield 1971). Participants were naïve to NIRS-BCIs, had normal or corrected-to-normal vision and had no known trauma-induced brain injuries, degenerative disorders, cardiovascular disorders, motor impairments, respiratory disorders, drug or alcohol-related conditions, psychiatric conditions or metabolic disorders. Participants were asked not to smoke or drink alcoholic or caffeinated beverages three hours prior to each data collection session. The study was conducted with informed consent and with ethics approval from the Holland Bloorview Kids Rehabilitation Hospital and the University of Toronto.

2.3.2 Instrumentation

NIRS data were collected using a multi-channel frequency-domain NIRS system with a sampling rate of 31.25 Hz (Imagent Functional Brain Imaging System from ISS Inc., Champaign, IL (ISS Inc. 2012)). The NIRS system was used to measure the blood oxygen content from the prefrontal cortex (PFC) (Ogata, Mukai, and Yagi 2007; Gao et al. 1990). Five laser diodes (emitting 690 nm and 830 nm light) and three photomultiplier tube detectors attached to a headband were used. The headband was centered on the participant’s forehead with reference to the nose, and was placed directly above the eyebrows, as illustrated in Figure 2.

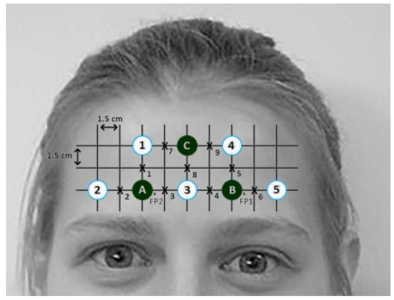

Figure 2. Experimental source and detector configuration. The solid circles represent detectors; the open circles represent light source pairs; the x's represent points of interrogation; and the starred areas represent the approximate FP1 and FP2 positions of the international 10-20 EEG system.

The sources and detectors were separated by a distance of 3 cm, which has been shown to reach the outer layer of the cerebral cortex (Bauernfeind et al. 2008; Haeussinger et al. 2011; E. Okada et al. 1997). The source-detector configuration allowed for the interrogation of nine discrete locations. A schematic diagram of the configuration and points of interrogation are illustrated in Figure 2.

2.3.3 Experimental protocol

Participants performed five sessions, one session each on five different days. Each session consisted of three data collection blocks. A schematic illustration of the study and block structure is shown in Figure 3. During each data collection block, the participant performed 24 task intervals. Each task was performed for 20 seconds, and was followed by a 17-second rest interval. Based on preliminary analysis, seventeen seconds allowed for an adequate amount of time for the haemodynamic activity from the task and ease-of-use selection to subside. In the literature, a variety of rest intervals have been used; including, twelve seconds (Sarah D Power, Kushki, and Chau 2012; L. C. Schudlo, Power, and Chau 2013; L. C. Schudlo and Chau 2014; Sarah Dianne Power and Chau 2013), fourteen seconds (Yoko Hoshi et al. 2011), fifteen seconds (S. M. Coyle, Ward, and Markham 2007; Herff, Heger, Fortmann, et al. 2013; Herff, Heger, Putze, et al. 2013; Meltem Izzetoglu et al. 2007), twenty seconds (Kelly Tai and Chau 2009; Naseer and Hong 2013a; Naseer and Hong 2013b; Naseer, Hong, and Hong 2014; Ogata, Mukai, and Yagi 2007), and thirty seconds (Ayaz et al. 2012).

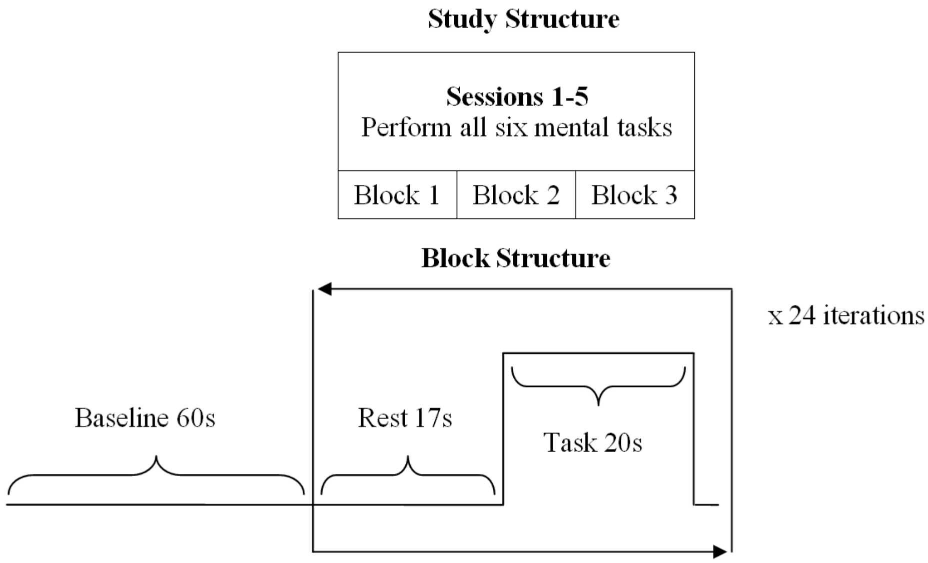

Figure 3. Study and block structure.

Two forms of neurofeedback were provided during all sessions: 1) a trapezoid topographic image showing the real-time changes in blood oxygenation levels over the PFC and 2) a ball that rose and fell with the average change over the entire interrogation area. The feedback was updated every 125 ms, and was calculated using cubic interpolation of the oxygenated haemoglobin (HbO) values at equally spaced intervals between the points of interrogation. The topographic image was 21 pixels in height with parallel sides 21 and 61 pixels in length, as in (L. C. Schudlo and Chau 2014). HbO was selected for the feedback since it has been cited to be more indicative of activity than deoxygenated haemoglobin (Hb) and total haemoglobin (tHb) (S. M. Coyle, Ward, and Markham 2007; S. Coyle et al. 2004). The red colour on the feedback represented an increase in haemodynamic activity, while the blue colour represented a decrease in haemodynamic activity. The goal of the neurofeedback was to provide participants with real-time information on changes in their haemodynamic activity when performing each of the tasks. Participants were informed that they should not stop performing the tasks; however, they could slightly modify the tasks, i.e. perform the tasks faster or slower, in order to try and achieve a more consistent change. In a study by Schudlo et al., it was found that 8 out of 10 participants adjusted their mental strategies when using feedback (L. C. Schudlo and Chau 2014). The user interface and haemodynamic feedback is shown in Figure 4.

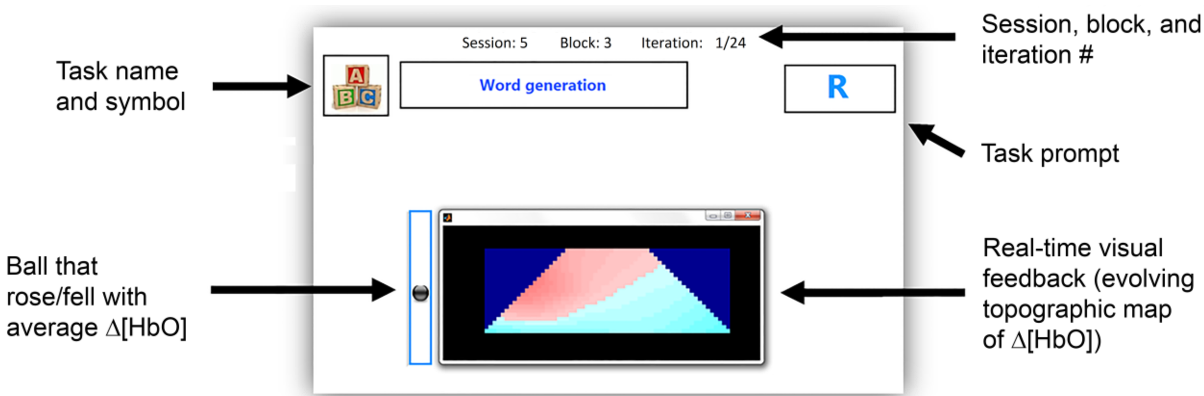

Figure 4. User interface and haemodynamic feedback.

Participants were asked to perform six different mental tasks, which are described in Table 1. The tasks were chosen based on their use in previous BCI studies or functional imaging studies.

Table 1. Six mental tasks performed by each participant.

Task	Description
Mental math (Math)	Participants were prompted with a math problem, and they were asked to repeatedly subtract a two digit number from a three digit number (Ang, Yu, and Guan 2012; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Ogata, Mukai, and Yagi 2007; Bauernfeind et al. 2008; Sarah D. Power, Kushki, and Chau 2012; Utsugi et al. 2007; L. C. Schudlo and Chau 2014).
Word generation (Words)	Participants were asked to think of as many words as possible that start with a specific letter (Ogata, Mukai, and Yagi 2007; Utsugi et al. 2007; Faress and Chau 2013).
Counting (Count)	Participants were asked to slowly count in their heads while relaxing (Naseer and Hong 2013a).
Happy thoughts (Happy)	Participants were asked to think about the details of a past event in their life that made them very happy (Kelly Tai and Chau 2009; Koshino et al. 2011).
Focusing (Focus)	Participants were asked to relax and focus on the feedback (K. Izzetoglu et al. 2011).
Rest (Rest)	Participants were asked to let their minds wander (Ang, Yu, and Guan 2012; Ayaz et al. 2007; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012; Naseer and Hong 2013a; L. C. Schudlo and Chau 2014).

The participants performed each of the six tasks four times per block in a random order. By the end of the fifth session, participants had performed each task 60 times. Immediately after performing each task and before the seventeen second rest, participants rated the task's perceived ease-of-use on a 5-point Likert-type scale, ranging from "very easy" to "very difficult" (Tedesco and Tullis 2006; Sauro and Dumas 2009).

At the end of the fifth session, participants were asked to choose their personalized tasks using the two user-centered selection methods, which will be described in section 2.3.4. The order in which the frameworks were presented was varied randomly.

2.3.4 Mental task selection frameworks

2.3.4.1 Prescribed

The state-of-the-art prescribed framework involves selecting the most commonly used set of tasks in NIRS-BCI literature for all participants. Specifically, mental math and rest were chosen as the two tasks (Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012; Ang, Yu, and Guan 2012; L. C. Schudlo and Chau 2014).

2.3.4.2 User-selected weighted slope scores (WS-US)

The user-selected weighted slope score (WS-US) framework involves each participant choosing their own two tasks based on a weighted slope score (WS-score) and the average perceived ease-of-use rating for each task, across all iterations. The WS-score, WS_i, for the i^{th} task, is defined as:

$$WS_i = \frac{\frac{1}{N}\sum_{j=1}^{N} m_{ij}}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(m_{ij} - \left[\frac{1}{N}\sum_{k=1}^{N} m_{ik}\right]\right)^2}} \quad (1)$$

where m_{ij} and m_{ik} are the slopes of least square line of best fit to the average haemodynamic activity ([HbO]) over time for the j^{th} or k^{th} iteration of the i^{th} task (Bingham and Fry 2010; Dupont 2009), and N is the number of times the task was performed (N = 60 in this study). The weighted slope score is thus the mean of all slopes for each iteration a task was performed divided by the corresponding standard deviation, and corresponds to the tendency for a task to consistently increase or decrease hemodynamic activity. Users were presented with the average perceived ease-of-use ratings and graphs of the best three tasks for increasing and decreasing their haemodynamic activity (based on the WS-score), and were asked to select one increasing and one decreasing task. Users were informed that tasks higher on the list resulted in more consistent and stronger changes in haemodynamic activity and that they should use this performance information in combination with their perceived ease-of-use data to choose their personalized tasks.

2.3.4.3 User-selected pair-wise accuracy rankings (PWAR-US)

The user-selected pair-wise accuracy rankings (PWAR-US) framework involves each participant choosing their own two tasks based on pair-wise accuracy rankings (PWAR) and the average perceived ease-of-use rating for each task, across all iterations. The accuracies were calculated after session five using one-iteration of six-fold cross-validation with eight selected features for each of the possible pair-wise combinations of the six mental tasks (6 choose 2 = 15). The pairs of tasks were then ranked from highest to lowest. The PWAR and average perceived ease-of-use ratings were then displayed to the user, and users where asked to select one pair of tasks from the list. Users were informed that tasks higher on the list resulted in better classification accuracies, and that they should use this performance information in combination with their perceived ease-of-use data to choose their personalized tasks.

2.3.4.4 Researcher-selected pair-wise accuracy rankings (PWAR-RS)

The researcher-selected pair-wise accuracy rankings (PWAR-RS) framework involves the researcher choosing the best set of tasks (highest ranked) for each participant based on PWAR. The PWAR were calculated as described in section 2.3.4.3.

2.3.5 Data processing

2.3.5.1 Filtering

To mitigate the effect of various physiological noise, the NIRS signal was digitally low-pass filtered in real-time using a third-order Chebyshev infinite impulse response (IIR) cascade filter with a pass-band from 0 to 0.1 Hz, a transition band from 0.1 to 0.5 Hz, a stop-band from 0.5 Hz onwards, and a pass band ripple of 0.1 dB (Sarah D. Power, Falk, and Chau 2010; Ayaz et al. 2009).

2.3.5.2 Calculating haemoglobin concentrations

After filtering the data, the changes in concentrations of HbO, Hb, and tHb were calculated using the modified Beer-Lambert's Law (S. M. Coyle, Ward, and Markham 2007; J. Wolpaw et al. 2000; Niels Birbaumer and Cohen 2007; Kelly Tai and Chau 2009). In this study, the constants used were r = 3cm, $\varepsilon_{690nm,Hb}$ = 2.1382 mM^{-1}cm^{-1} (Cope 1991), $\varepsilon_{830nm,Hb}$ = 0.7804 mM^{-1} cm^{-1} (Cope 1991), $\varepsilon_{690nm,HbO}$ = 0.3123 mM^{-1} cm^{-1}(Cope 1991), $\varepsilon_{830nm,HbO}$ = 1.0507 mM^{-1} cm^{-1} (Cope 1991), DPF_{690nm} = 6.51 (A Duncan et al. 1995), and DPF_{830nm} = 5.86 (A Duncan et al. 1995).

2.3.5.3 Feature extraction

Features were extracted over four time-windows (0-5s, 0-10s, 0-15s, and 0-20s). Features included the temporal changes in the three chromophores (Hb, HbO, and tHb) at each of the 9 points of interrogation (108 features) and the spatial features of the zero to fourth order discrete orthogonal Chebyshev image moments (180 features), as proposed in (L. C. Schudlo, Power, and Chau 2013). The temporal feature extraction involved normalizing each task interval and then determining the least square line of best fit slope over the different time-windows. To derive the spatial features, topographic images for Δ[HbO], Δ[Hb], and Δ[tHb] were generated by cubic interpolation of the hemoglobin concentration values between locations of empirical integration. The images were normalized, and the zero to fourth order discrete orthogonal Chebyshev polynomial image moments were extracted from the dynamic topograms. Chebyshev polynomials were calculated using equation 9, equation 12, and Table 2 from (Zhu et al. 2010). For more information on the feature extraction methods, please refer to (L. C. Schudlo, Power, and Chau 2013). A block diagram summarizing the feature extraction methods is shown in Figure 5.

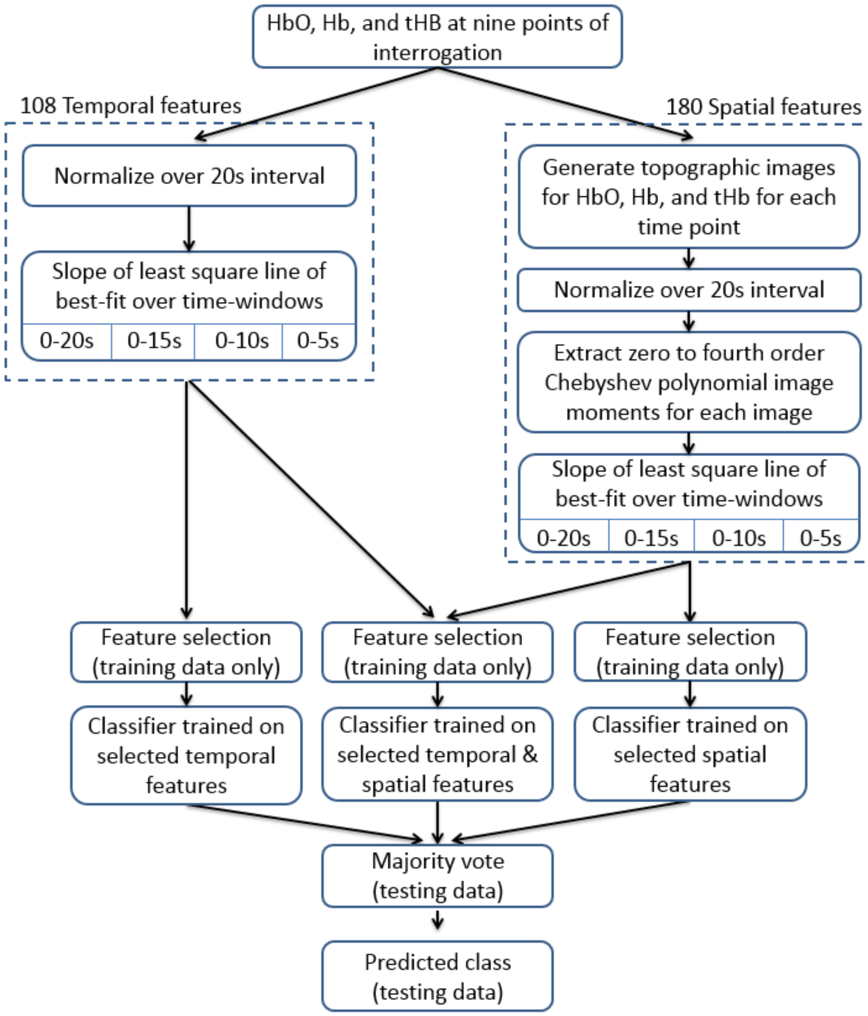

Figure 5. Block diagram of feature extraction and classification methods.

2.3.5.4 Cross-validation, feature selection, and pattern classification

Data collected across all five sessions were pooled together, and accuracies were determined using 10-fold cross-validation (Refaeilzadeh, Tang, and Liu 2009). For each iteration of 10-fold cross-validation, the data were randomly separated into 10 equal sized portions (folds). Ten classification accuracies were calculated by iteratively using each fold as testing data and the remaining folds as training data. Only training data were used for feature selection and classifier training, and only the testing data were used to estimate the classification accuracies. Finally, all classification accuracies were averaged to estimate the overall accuracy.

A fast correlation based filter (FCBF) was used to select a subset of up to eight features from the total feature set for classifier training (Yu and Liu 2003; Koelstra et al. 2010; Chanel, Ansari-

Asl, and Pun 2007). A judicious subset of features has been shown to lead to smaller classification errors (Ang, Yu, and Guan 2012). The upper limit of features was based on preliminary data and past work (Sarah D Power, Kushki, and Chau 2012; L. C. Schudlo and Chau 2014) and aimed to maintain an adequate ratio of training samples to features, i.e. avoiding the "curse of dimensionality" (Hastie, Tibshirani, and Friedman 2009).

An ensemble of classifiers was invoked for each participant to differentiate between task-induced changes in the haemodynamic response (Polikar 2006; L. C. Schudlo, Power, and Chau 2013). In particular, for each participant, three ensemble classifiers with ten members of linear discriminant classifiers were trained: one exclusively with temporal features, a second exclusively with spatial features, and a third using a combination of temporal and spatial features. The majority vote (Polikar 2006) of the classifiers was used as the class prediction. A block diagram summarizing the classification methods is shown in Figure 5.

2.3.5.5 Pseudo-online BCI simulation

A secondary pseudo-online analysis (using sessions 1 and 2 for training and sessions 3 - 5 for testing) was conducted on the researcher-selected PWAR tasks and prescribed tasks to simulate online BCI performance and verify the benefit of researcher-selected tasks with respect to accuracy. One-iteration of six-fold cross-validation was performed on the data collected in sessions 1 and 2, and the best researcher-selected tasks were chosen based on PWAR. The data from sessions 1 and 2 (24 samples per class) were then used for feature selection and to train the ensemble of classifiers. The classifiers were then used to predict the classes of the data from sessions 3 to 5 (36 samples per class). The same feature extraction (temporal and spatial), feature selection (FCBF), and classification (majority vote) methods described in sections 2.3.5.3 and 2.3.5.4 were invoked.

2.3.6 Data analyses

For all statistical tests, normality of the data was confirmed using the Shapiro-Wilk Normality test.

2.3.6.1 Accuracy comparison

A two-tailed paired Student's *t*-test for two dependent means was used to compare the accuracies attained.

2.3.6.2 Perceived ease-of-use comparison

To compare the perceived ease-of-use, the average total subjective perceived ease-of-use ratings for each set of tasks was determined, and a two-tailed paired Student’s t-tests for two dependent means was run.

2.3.6.3 Computational time comparison

To compare the computational time of each framework, a time analysis was performed on a desktop computer with a Intel Core2 Quad Q8300, 4Mb cache, 2.50 GHz processor, 2 Gb of RAM, running Windows XP.

2.3.6.4 User preference of selection methods

To determine which of the two user-centered choosing methods participants preferred, a questionnaire was administered at the end of the fifth session. Specifically, participants evaluated the following subjective statement: “Did you prefer choosing your tasks using the 'best tasks to increase/decrease activity' page or the 'suggested task pairs' page? Please explain why”.

2.3.6.5 Data collection requirements

To determine how much data should be collected before the task selection is made, an analysis of the cumulative WS-score and PWAR over time was conducted. The WS-score and accuracies were calculated after every block, and the absolute value of the regression slope between two adjacent points was determined. A sufficient amount of data was deemed to have been collected when two conditions held true: 1) two consecutive slopes had a value less than 0.05 and 2) the slope between the first and third point was also less than 0.05. Offline accuracies were calculated using five iterations of ten-fold cross validation.

2.3.6.6 Analysis of selected features

We examined the average number of features selected and the frequency at which each chromophore (Hb, HbO, and tHb) and time-window (0-5s, 0-10s, 0-15s, and 0-20s) was selected. The analysis was conducted over all ten participants for each of the three classifiers (temporal, spatial, and temporal combined with spatial), using all data collected in sessions 1 to 5, and for all four frameworks.

2.4 Results

2.4.1 Tasks chosen

The tasks chosen using each of the frameworks is shown in Table 2. It is noted that three out of ten participants (P5, P9, and P10) chose the exact same tasks using both user-centered selection methods. It is also noted that only one participant (P3) selected the same tasks using PWAR as the tasks chosen by the researcher.

Table 2. Tasks chosen using each of the mental task selection frameworks: prescribed, user-selected using weighted slope scores (WS-US), user-selected using pair-wise accuracy rankings (PWAR-US), and researcher-selected using pair-wise accuracy rankings (PWAR-RS). Legend: math - mental subtraction, words - word generation, count - counting slowly, happy - thinking of happy thoughts, focus - focusing on the feedback, and rest - letting your mind wander. For more information on the mental tasks, please refer to Table 1.

Participant	Prescribed	WS-US	PWAR-US	PWAR-RS
1	Math & Rest	Happy & Rest	Words & Focus	Math & Focus
2	Math & Rest	Happy & Count	Words & Focus	Words & Rest
3	Math & Rest	Words & Rest	Happy & Rest	Happy & Rest
4	Math & Rest	Math & Words	Words & Count	Math & Happy
5	Math & Rest	Words & Count	Words & Count	Math & Focus
6	Math & Rest	Words & Count	Words & Rest	Happy & Focus
7	Math & Rest	Focus & Count	Focus & Rest	Math & Rest
8	Math & Rest	Words & Count	Math & Focus	Happy & Count
9	Math & Rest	Focus & Count	Focus & Count	Happy & Focus
10	Math & Rest	Words & Count	Words & Count	Happy & Focus

2.4.2 Researcher-selected tasks result in significantly higher accuracies

The average offline cross-validation accuracies and standard deviations for prescribed tasks, WS-US tasks, PWAR-US tasks, and PWAR-RS tasks were: 65.3 ± 4.5%, 65.7 ± 8.3%, 68.9 ± 9.6%, and 76.6 ± 8.2%, respectively, as shown in Figure 6.

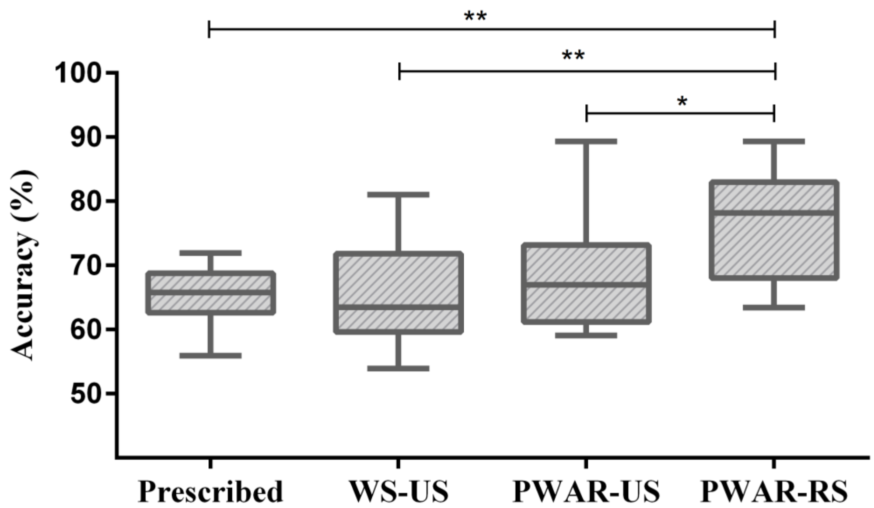

Figure 6. Box plots of task accuracies for each of the four mental task frameworks. Legend: Prescribed = prescribed tasks (mental math and rest), WS-US = user-selected tasks using weighted slope scores, PWAR-US = user-selected tasks using pair-wise accuracy rankings, PWAR-RS = researcher-selected tasks using pair-wise accuracy rankings, * = $p < 0.05$, and ** = $p < 0.01$.

There was no significant difference in the offline accuracies between the WS-US tasks, the PWAR-US tasks, and the prescribed tasks. However, the accuracies achieved using the PWAR-RS tasks were significantly higher than the prescribed tasks ($t = 3.91$, $p = 0.0036$), WS-US tasks ($t = 4.351$, $p = 0.0018$), and PWAR-US tasks ($t = 2.63$, $p = 0.0273$).

For the pseudo-online BCI simulation, the average accuracy for the PWAR-RS tasks was 72.1 ± 10%, while the average accuracy for the prescribed tasks was 62.6 ± 7%. The two-tailed paired student's *t*-test showed that the personalized researcher-selected tasks resulted in significantly greater accuracies than the prescribed tasks ($t = 2.66$, $p = 0.0257$).

2.4.3 User-selected tasks result in significantly higher perceived ease-of-use

The average perceived ease-of-use and standard deviations of the prescribed tasks, WS-US tasks, PWAR-US tasks, and PWAR-RS tasks were: 3.44 ± 1.06, 4.19 ± 0.439, 4.3 ± 0.485, and 3.50 ± 0.899, respectively, as shown in Figure 7.

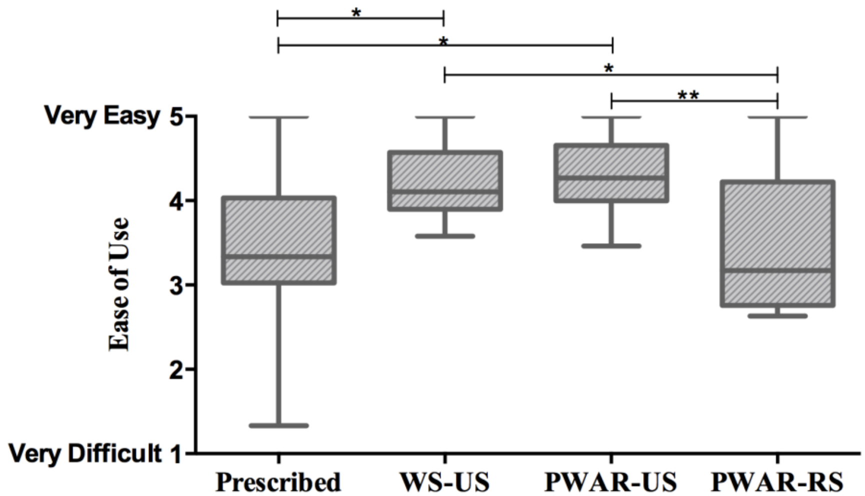

Figure 7. Box plots of task perceived ease-of-use for each of the four mental task frameworks. Legend: Prescribed = prescribed tasks (mental math and rest), WS-US = user-selected tasks using weighted slope scores, PWAR-US = user-selected tasks using pair-wise accuracy rankings, PWAR-RS = researcher-selected tasks using pair-wise accuracy rankings, * = *p* < 0.05, and ** = *p* < 0.01.

Users found tasks chosen using the two user-centered selection methods to be significantly easier to use than both the prescribed tasks and the researcher-selected personalized tasks. Specifically, the prescribed tasks were significantly harder to use than WS-US tasks (t = 2.778, p = *0.0215*) and PWAR-US tasks (t = 3.101, p = *0.0127*). Moreover, PWAR-RS tasks were significantly harder to use than WS-US tasks (t = 2.982, p = *0.0154*) and PWAR-US tasks (t = 3.969, p = *0.0033*).

2.4.4 PWAR have longest computational time

Obviously, prescribed tasks require no calculations prior to task selection. The computational time of all six WS-scores only took 30 seconds. The computational time for all 15 pair-wise combinations of tasks was approximately 6 minutes. It is noted that this time will vary depending on the speed of the computer and the implementation of the algorithm.

2.4.5 Users prefer choosing tasks using WS-score rather than PWAR

When comparing user preference between the two user-centered selection methods, it was found that 8 out of 10 participants preferred choosing their personalized tasks using WS-scores compared to PWAR.

2.4.6 WS-scores require largest data set before task selection

Obviously, prescribed tasks require no data to be collected prior to task selection. On the other hand, it was determined that on average 35 data points per class need to be collected before the WS-scores should be used to guide the selection of personalized tasks. However, only 23 data points per class need to be collected before using PWAR to inform the selection of personalized mental tasks. Given the current study structure, this would be equivalent to three and two sessions of data collection for WS-Scores and PWAR, respectively.

2.4.7 Feature selection analysis

On average 4.9 features were selected across all three classifiers and all four frameworks, as shown in Figure 8A. It appears that more features were selected for the combined temporal and spatial classifier than the temporal-only and spatial-only classifiers for all four frameworks. All three chromophores (Hb, HbO, tHb) and all four time-windows (0-5s, 0-10s, 0-15s, and 0-20s) were frequently chosen during features selection for each participant. Overall, features from the chromophore Hb and the 0-17s time-window were chosen most often, as shown in Figure 8B and Figure 8C, respectively.

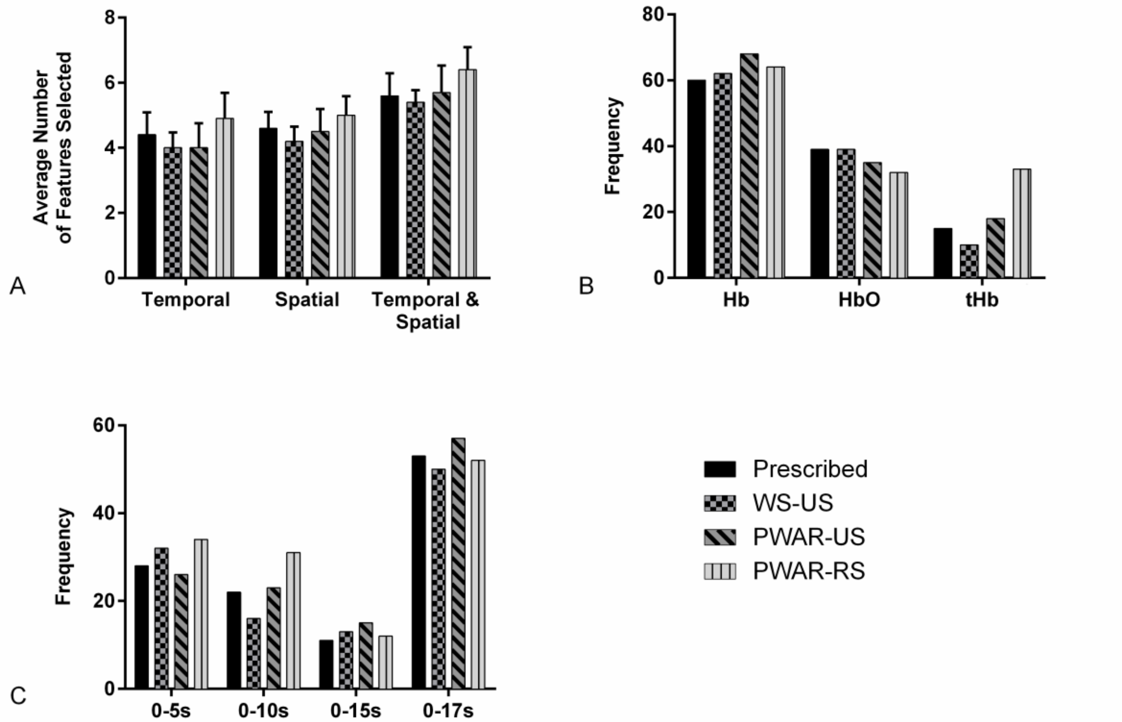

Figure 8. (A) Average number of features selected. (B) Frequency of occurrence of each chromophore (Hb, HbO, and tHb). (C) Frequency of occurrence of each time-window (0-5s, 0-10s, 0-15s, and 0-20s).

2.5 Discussion

In order to provide a quick guide for researchers to choose a mental task framework for future NIRS-BCI studies, Table 3 summarizes each framework's performance in terms of the five metrics.

Table 3. A summary of the performance of each of the four frameworks with respect to accuracy, perceived ease-of-use, computational time, preference, and amount of data that needs to be collected prior to use. Legend: Prescribed = prescribed tasks (mental math and rest), WS-US = user-selected tasks using weighted slope scores, PWAR-US = user-selected tasks using pair-wise accuracy rankings, and PWAR-RS = researcher-selected tasks using pair-wise accuracy rankings.

Metric	Prescribed	WS-US	PWAR-US	PWAR-RC
Accuracy	65.3 ± 4.5%	65.7 ± 8.3%	68.9 ± 9.6%	76.6 ± 8.2%
Ease-of-use (1 = very difficult; 5 = very easy)	3.44 ± 1.06	4.19 ± 0.44	4.30 ± 0.49	3.50 ± 0.90
Computational time	~ 0 seconds	~ 30 seconds	~ 6 minutes	~ 6 minutes
Preference	Not assessed	Preferred by 8/10	Preferred by 2/10	Not assessed
Amount of data before choosing	0 data points	35 data points	23 data points	23 data points

2.5.1 Benefits of personalized tasks over prescribed tasks

Since none of the frameworks outperforms on all metrics, it is imperative to look at the relative importance of each metric. Given that accuracy and usability are two of the most important metrics in BCIs (J. Wolpaw et al. 2000; Bos, Poel, and Nijholt 2011; Tan and Nijholt 2010; Holz et al. 2013), it is clear that personalized tasks are superior to prescribed tasks. Specifically, compared to prescribed tasks, researcher-selected tasks resulted in higher accuracies and user-selected tasks resulted in higher perceived ease-of-use. These results combined with the large inter-subject variability, support the value of personalization. Overall, these findings are in line with previous EEG-BCI and MRI-BCI studies that have also found a large inter-subject variability and an association between researcher-selected personalized mental tasks and an increase in accuracy (Sorger et al. 2009; Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012). Moreover, the benefits of personalization are in line with findings in other areas of research such as in education. Specifically, personalization has been shown to increase learning, motivation, and depth of engagement in education (Cordova and Lepper 1996).

2.5.2 Best personalized framework

Each of the three personalized selection methods explored in this work appears to have different benefits. Specifically, the PWAR-RS framework maximizes accuracy and minimizes the amount of data that needs to be collected. On the other hand, the PWAR-US framework maximizes usability and minimizes the amount of data that needs to be collected. Finally, the WS-US framework maximizes usability, maximizes user satisfaction, and minimizes computational time. The authors suggest that the choice of personalized BCI framework be based on the metric deemed most important for the application at hand.

2.5.2.1 Accuracy and perceived ease-of-use

As mentioned in section 2.5.1, both accuracy and usability are often cited to be two of the most important metrics of BCIs (J. Wolpaw et al. 2000; Bos, Poel, and Nijholt 2011; Tan and Nijholt 2010; Holz et al. 2013). Which of the two is more important has not been robustly established, and may even depend on the individual use-case. Further research is required in order to establish whether usability or accuracy is deemed to be the single most important BCI metric or whether this is a multi-criteria decision. Moreover, the long term implications of both measures still need to be explored in future research. For example, it is possible that increased usability could translate to increased accuracy over time.

2.5.2.2 Computational time, user preference, and training data

Overall, both PWAR and WS-scores have advantages and limitations when it comes to computational time, user preference, and data required for task selection.

PWAR task selection was computationally more demanding than selection by WS-scores. Moreover, the computational time of PWAR increases quadratically with the number of mental tasks, specifically, the number of pair-wise combinations is $(n)*(n-1)/2$, where n is the number of tasks. For example, with 6 tasks, there are only 15 pair-wise combinations. However, with 10 tasks there would be 45 pair-wise combinations. In comparison, the computational time of WS-scores increases linearly with the number of tasks. The higher computational complexity of PWAR could be a major limitation as users are often eager to choose their tasks and do not want to wait for data to be processed.

Overall, users appeared to prefer using WS-scores. Based on user feedback, this is likely due to the higher transparency of this method. Users were able to understand that they were choosing one task that tended to increase the haemodynamic activity and one task that tended to decrease the haemodynamic activity. Thus, there was clarity in terms of what to look for in the neurofeedback. On the other hand, PWAR is more of a black-box approach from a user's perspective. Users are told that the computer can tell some pairs of tasks apart better than others, but they aren't given a reason as to why or how. Additionally, when using WS-scores, participants only have to deal with six items on the screen, three tasks for increasing and three tasks for decreasing haemodynamic activity; whereas, for PWAR, there were 15 possible pair-wise combinations displayed. The increased amount of information may have been more difficult for users to process.

Finally, the fewer data samples required before using PWAR for task selection could be considered a major advantage. Specifically, fewer training sessions would be needed before the user can start to utilize the BCI. In the research context, time is usually a major factor and the shorter the study the more feasible its completion.

2.5.3 User accuracy/usability trade-off

When looking at the tasks chosen by users using PWAR, it is evident that perceived ease-of-use and preferences play a large role in task selection. Despite being presented with the same PWAR, only one participant (P3) selected the same tasks as those chosen by the researcher.

This speaks to the importance that users place on ease-of-use, and gives insight into the way users traverse the accuracy/usability trade-off. It is noted that all users were informed that by not choosing the highest performing set of tasks, they would be sacrificing accuracy for ease-of-use. The lowest PWAR chosen was a rank of 7 out of the 15 combinations. Specifically, P6, P7, and P8 all chose the seventh ranked pair of tasks. It is also noted that on average, the fifth set of tasks from the PWAR was chosen.

The accuracy/usability trade-off has not received much attention in research; however, similar trade-offs have been explored in other areas of research. One of the most well researched areas is the speed/accuracy trade-off made by individuals on a daily basis (Wickelgren 1977; Bogacz et al. 2010). Further exploration of the accuracy/usability trade-off should be conducted.

2.5.4 Significance

To the best of our knowledge, this is the first NIRS-BCI study to explore personalized mental tasks. Researcher-selected tasks have previously been explored in MRI-BCI (Sorger et al. 2009) and EEG-BCI studies (Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012); however, their application to NIRS-BCI's has not been explored. Since different BCI access modalities differ in their methods of measurement and in terms of their measured responses, the results from one modality cannot be extended to others. Moreover, to the best of our knowledge, this is the first BCI study that has allowed users to choose their own user-selected personalized tasks based on task ease-of-use and performance. Since this work, we have used personalized mental tasks in two other NIRS-BCI studies (Weyand, Takehara-Nishiuchi, and Chau 2015b; Weyand et al. 2015).

2.5.5 Limitations and future work

This study was conducted under ideal environmental conditions (quiet and dimly-lit room), which may not be indicative of most real-world settings. Further research should be conducted to assess the effect of environmental conditions. Secondly, this study was conducted offline; further investigation is required to ensure that these results translate to online BCIs. Thirdly, this study was conducted over a relatively short period of time, spanning only five sessions. It is possible that accuracy and usability measures change over a longer period of time. Fourthly, this study provided users with feedback to reinforce task performance, which may have resulted in divided attention between performing the task and observing the feedback (Quinlan and Dyson 2008). It has been shown in the literature that dual-task paradigms often result in similar activations of brain regions as if the tasks were performed separately; however, possibly with a lower overall activation due to the divided attention (Newman, Keller, and Just 2007). It is noted that lower task activation is not necessarily a detrimental limitation so long as the changes in brain activations can still be classified. Finally, this study has only investigated four task selection frameworks. Variations of these frameworks and alternative frameworks could also be explored in future research.

2.6 Conclusions

This study explored four NIRS-BCI mental task selection frameworks. It was shown that no single framework maximized all metrics. However, in general our findings show that personalized rather than prescribed tasks are associated with increased accuracy and perceived ease-of-use. Specifically, researcher-selected personalized tasks yielded the highest accuracies, while user-selected personalized tasks resulted in the highest perceived ease-of-use. When comparing the personalized task selection methods, it was concluded that PWAR minimized the amount of data that needs to be collected. On the other hand, WS-scores maximized user satisfaction, and minimized computational time. Overall, this research provides an incentive for the further exploration of personalized mental tasks in future NIRS-BCI studies.

Chapter 3: Correlates of User Characteristics with Accuracy and Moving Beyond Binary Classification

The entirety of this chapter is reproduced from the article "Correlates of near-infrared spectroscopy brain-computer interface accuracy in a multi-class personalization framework". This manuscript has been published in Frontiers in Human Neuroscience.

3.1 Abstract

Brain-computer interfaces (BCIs) provide individuals with a means of interacting with a computer using only neural activity. To date, the majority of near-infrared spectroscopy (NIRS) BCIs have used prescribed tasks to achieve binary control. The goals of this study were to evaluate the possibility of using a personalized approach to establish control of a 2-, 3-, 4-, and 5-class NIRS-BCI, and to explore how various user characteristics correlate to accuracy. Ten able-bodied participants were recruited for five data collection sessions. Participants performed six mental tasks, and a personalized approach was used to select each individual's best discriminating subset of tasks. The average offline cross-validation accuracies achieved were 78%, 61%, 47%, and 37% for the 2-, 3-, 4-, and 5-class problems, respectively. Most notably, all participants exceeded an accuracy of 70% for the 2-class problem, and two participants exceeded an accuracy of 70% for the 3-class problem. Additionally, accuracy was found to be strongly positively correlated (Pearson's) with perceived ease of session (ρ = 0.653), ease of concentration (ρ = 0.634), and enjoyment (ρ = 0.550), but strongly negatively correlated with verbal IQ (ρ = -0.749).

3.2 Introduction

3.2.1 Near-infrared spectroscopy brain-computer interface

Brain-computer interfaces (BCIs) can be used as an access pathway for individuals with severe motor impairments as they require only brain activations and no muscular control (BZ Allison, Wolpaw, and Wolpaw 2007). Near-infrared spectroscopy (NIRS) has recently gained attention as a BCI access modality due to its non-invasive extraction methods, gel-less donning, and robustness to electrical noise (Ferrari, Mottola, and Quaresima 2004; Scholkmann et al. 2014). In general, NIRS can be used to detect changes in the amount of oxygen in neuronal blood, which reflect changes in brain activation (Scholkmann et al. 2014; Boas et al. 2014; Strait and Scheutz 2014). A computer can be trained to discriminate between mental tasks based on changes of the hemodynamic response resulting from the task being performed.

Currently, most hemodynamic BCIs use two prescribed tasks to achieve binary control of a computer. When using binary BCIs for communication, these tasks can be mapped to a “scroll and select” or “yes and no” output (Naito et al. 2007). A handful of studies have been conducted on NIRS-BCIs over the prefrontal cortex (PFC), achieving average accuracies ranging from around 60% to 80% (Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Naseer and Hong 2013a).

Multiclass BCIs (beyond binary) have the potential to provide users with more outputs, thereby increasing the rate of communication (Shin et al. 2013). However, as the number of classes increases, so will the difficulty in discriminating between classes. To date, limited research on multi-class NIRS-BCIs has been conducted. To the best of our knowledge, three studies have explicitly explored multi-class NIRS-BCIs over the PFC that could potentially be used for active computer control, namely, (Herff, Heger, Fortmann, et al. 2013), (Hirshfield et al. 2009), and (Sarah D Power, Kushki, and Chau 2012). Herff et al. classified mental workload states, using the n-back task with an average 3-class and 4-class accuracy of 50.3% and 44.5%, respectively (Herff, Heger, Fortmann, et al. 2013). Hirshfield et al. discriminated between different levels of mental workloads, achieving an average 3-class accuracy of 54% (Hirshfield et al. 2009). Finally, Power et al. were able to distinguish between mental math, mental singing and rest with an average accuracy of 56.2% (Sarah D Power, Kushki, and Chau 2012). A second study by Herff et al. also explored differentiating mental arithmetic, word generation, and mental rotation. Although the accuracies for a 3-class problem were not explicitly stated, the authors indicated that the 3-class accuracies were greater than chance (Herff, Heger, Putze, et al. 2013). Overall, these studies demonstrate proof-of-concept for a multi-class NIRS-BCI over the PFC, but are not at the level that is required for effective BCI use. It appears that none of the participants in these studies exceeded the 70% threshold, often cited as required for BCI control (Andrea Kübler, Neumann, et al. 2001).

3.2.2 Personalized tasks

One potential method for improving the classification accuracies achievable in an NIRS-BCI is the use of researcher-selected personalized mental tasks, an approach whereby a user tries a variety of tasks and subsequently a sub-set of tasks that are most suitable for that user are selected by the researcher. Task selection is usually based on the discriminating power of the tasks. To date, personalized mental tasks have been explored in a 2-class offline NIRS-BCI study (Weyand, Takehara-Nishiuchi, and Chau 2015c), as well as in a 2-class magnetic resonance imaging (MRI) BCI study (Sorger et al. 2009) and in 2–class (Palaniappan 2006), 3-class (Chai et al. 2012), and 4-class (Dobrea and Dobrea 2009) electroencephalography (EEG) BCI studies. Overall, these studies conclude that there is significant inter-subject variability in brain activation elicited by the same mental tasks and cognitive processes, and as a result, the tasks that are most effective for controlling a BCI vary among users. Therefore, there is potential to improve classification accuracies by choosing the most discriminating tasks for each user (Sorger et al. 2009; Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012; Weyand, Takehara-Nishiuchi, and Chau 2015c). To the best of our knowledge, to date, personalized tasks have not been explored in an NIRS-BCI beyond the binary paradigm.

3.2.3 Correlation between BCI accuracy and user characteristics

Another sparsely explored area in literature is the prediction of BCI accuracy based on user characteristics, such as: demographic traits, IQ, and state of mind. Determining the correlation between user characteristics and performance may help to reduce some of the large inter-subject variability in classification accuracies, steer future BCI development, and provide additional measures for selecting user-specific tasks.

To date, limited accuracy-user correlation research has been conducted in the field of NIRS-BCIs. However, several studies have examined the inter- and intra-subject correlations between accuracy and characteristics of able-bodied participants using various EEG-based BCIs. Studies have reported increased accuracy to be correlated with: increasing self-reported task enjoyment (*Pearson's* $\rho = 0.3$, $p < 0.1$) (E V C Friedrich, Scherer, and Neuper 2013), increasing challenge (*Spearman's* $\rho = 0.8$, $p < 0.01$) (Kleih et al. 2011), decreasing sleep (*Mann-Whitney test,* $p < 0.05$) (Guger et al. 2009), increasing mood (*multiple regression* $b = 0.498$, $p < 0.05$) (Nijboer et al. 2008), increasing mastery confidence (*multiple regression* $b = 0.578$, $p < 0.05$) (Nijboer et al. 2008), and increasing attention (*Spearman's* $\rho = 0.5$, $p = 0.02$) (Hammer et al. 2012). Conflicting trends have been reported on the association of accuracy with fear of incompetence, i.e., anxiety of failing the task. Studies have noted increased accuracy with increasing (*Spearman's* $\rho = 0.37$, $p < 0.05$) (Kleih et al. 2011), and decreasing (*multiple* *regression b* = -0.616, $p < 0.05$) (Nijboer et al. 2008) fear of incompetence, both when visual feedback was provided. However, in the presence of auditory feedback accuracy and fear increased together (*multiple regression b* = 0.47, $p < 0.05$) (Nijboer et al. 2008).

Limited research has also been conducted on the correlation of accuracy with user demographics. Randolph et al. documented a positive relationship between age and control signal strength (*multiple linear regression*, $p = 0.013$) (Randolph, Karmakar, and Jackson 2006). Meanwhile, Allison et al. observed that older subjects and male subjects tended to perform worse; however, it is noted these trends were not significant (Brendan Allison et al. 2010).

In contrast to the above, several researchers found no correlations between accuracy and user characteristics; for example, in Guger et al., gender, education, work duration, and cigarette and coffee consumption were not statistically related to accuracy (Guger et al. 2009), and in Hammer et al., intelligence, mood, motivation, or learning abilities were not correlated with accuracy (Hammer et al. 2012).

In addition to the studies on able-bodied participants, Nijboer et al. performed an intra-subject correlation analysis on six individuals with amyotrophic lateral sclerosis to explore the effect of quality of life, depression, mood, mastery confidence, incompetence fear, interest, and challenge on performance over time (across sessions). They found that BCI performance was positively related to mastery (*Spearman's* ρ = +0.805) in one participant, positively related to challenge (*Spearman's* ρ = +0.733) in another, and negatively related to incompetence fear (*Spearman's* ρ = -0.824) in a third. No other correlations were found in the remaining three participants (Nijboer, Birbaumer, and Kübler 2010).

Although, to the best of our knowledge, no studies have explored correlations of NIRS-BCI performance with respect to user characteristics, the variety of correlations reported in EEG-BCI literature, along with the known dependencies of neural oxygenation levels on gender (F. Okada et al. 1993), handedness (F. Okada et al. 1993), age (Schroeter et al. 2003; Kwee and Nakada 2003), and IQ (Graham et al. 2010), suggest that such an investigation is warranted.

3.2.4 Objectives

The first objective of this study was to use a personalized mental task approach to determine the accuracies achievable for a 2-, 3-, 4-, and 5-class NIRS-BCI. The second objective was to ascertain the strength of the correlations between the accuracy achieved over five sessions by each participant and his or her personal characteristics, specifically, gender, handedness, age, verbal IQ score, average self-reported ease of session, average self-reported session enjoyment, average self-reported user tiredness, and average self-reported ease of concentration.

3.3 Methods

It is noted that the data collected during this study were also analyzed to compare 2-class prescribed and personalized NIRS-BCI frameworks offline. For more information on this work, please refer to (Weyand, Takehara-Nishiuchi, and Chau 2015c).

3.3.1 Participants

Ten able bodied participants (4 male, 6 female) were recruited from the staff and students at Holland Bloorview Kids Rehabilitation Hospital (Toronto, Canada). Signed consent was obtained from all participants and the study was approved by the ethics departments at Holland Bloorview Kids Rehabilitation Hospital and the University of Toronto. All participants were self-selected and naïve to NIRS-BCIs.

3.3.2 Criteria

Participants had normal or corrected-to-normal vision, were not receiving psychoactive medication such as anti-depressants or analgesics, and did not have any health conditions that may affect the measurement of or one's ability to perform mental tasks, including, but not limited to: degenerative disorders, cardiovascular disorders, metabolic disorders, trauma-induced brain injury, respiratory conditions, drug and alcohol-related conditions, and psychiatric disorders. Lastly, participants had to communicate in English, refrain from smoking, and avoid drinking alcohol or caffeinated beverages three hours prior to data collection.

3.3.3 Instrumentation

The Imagent Functional Brain Imaging System from ISS Inc., Champaign, IL (ISS Inc. 2012) was used to collect the NIRS data at a sampling rate of 31.25 Hz. Three photomultiplier tube detectors and five laser diodes (emitting 690 and 830 nm light) were arranged in a trapezoid, as shown in Figure 9. The trapezoid configuration allowed for discrete signal extraction at nine points of interrogation, located between each detector and diode that is separated by a distance of 3 cm (Naito et al. 2007). A headband made of rubber polymer (3M 9900 series) was used to position the detectors and light sources. All detectors and diodes were held in place by opaque fabric pockets that provided shielding from ambient light, and minimized detector and diode motion, while maximizing contact with the head. The headband was positioned above the eyebrows and centered with respect to the nose. Additionally, an accelerometer attached to the headband was used to collect information on head movement.

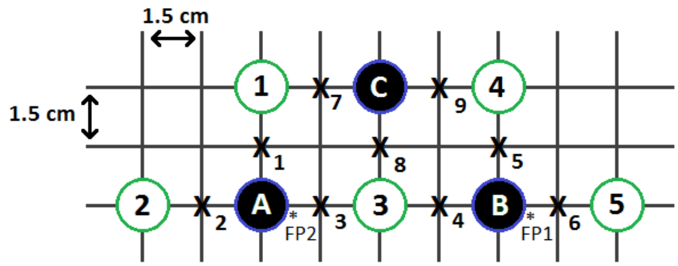

Figure 9. Experimental source and detector configuration. The solid circles represent detectors; the open circles represent light source pairs; the x's represent points of interrogation (channels); and the starred areas represent the approximate FP1 and FP2 positions of the international 10-20 EEG system.

3.3.4 Experimental protocol

3.3.4.1 Study structure and user interface

Participants took part in five data collection sessions on five separate days. Each session consisted of three data collection blocks. During each block, a baseline of 30 seconds was collected followed by 24 task intervals. The task intervals consisted of a task being performed for 20 seconds followed by a 17 second rest period. All six tasks were performed four times in a random order. Participants were asked to remain still during the task intervals.

Each task was performed 60 times by each participant (4 repetitions/block x 3 blocks/session x 5 sessions = 60).

Participants were provided with two forms of feedback: 1) A real-time trapezoidal topographic image that corresponded to the hemodynamic changes over the entire interrogation area, and 2) A ball which rose and fell depending on the average activation over the trapezoid. The goal of the neurofeedback was to provide participants with real-time information about changes in their haemodynamic activity when performing each of the tasks. Participants were informed that they should not stop performing the tasks; however, they could slightly modify the tasks, i.e. perform the tasks faster or slower, in order to try and achieve a more consistent change in the feedback. In a study by Schudlo and Chau, it was found that 8 out of 10 participants adjusted their mental strategies when using feedback (L. C. Schudlo and Chau 2014).The feedback was updated every 125 ms, and was calculated using cubic interpolation of the oxygenated haemoglobin (HbO) values at equally spaced intervals between the points of interrogation. The topographic image was 21 pixels in height with parallel sides 21 and 61 pixels in length, as in (L. C. Schudlo and Chau 2014; Weyand et al. 2015). The red colour on the feedback represented an increase in haemodynamic activity, while the blue colour represented a decrease in haemodynamic activity. The user interface, including the two types of feedback, is shown in Figure 10.

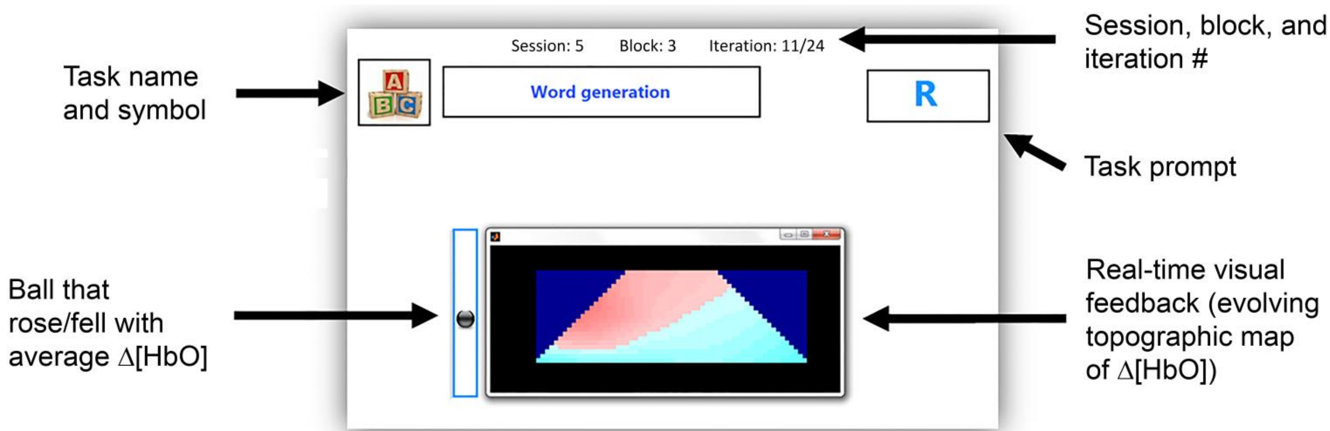

Figure 10. User interface for all sessions. The task name and symbol shows which of the six tasks the user should perform, i.e. mental math, word generation, happy thoughts, relaxing with focus, relaxing with slow counting, or unconstrained rest.

3.3.4.2 Mental tasks

In this study, we explored six mental tasks, selected on the basis of past literature indicating their suitability for NIRS-BCI control. Each of the six mental tasks is described in Table 4.

Table 4. Mental tasks and descriptions.

Mental Task	Description
Mental math (MM)	Users continuously subtract a randomly generated two-digit number from a randomly generated three-digit number. For example, given the equation "593-11", users would think "593-11 = 582; 582-11 = 571; 571-11 = 560…" (Ang, Yu, and Guan 2012; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Ogata, Mukai, and Yagi 2007; Bauernfeind et al. 2008; Sarah D. Power, Kushki, and Chau 2012; Utsugi et al. 2007; L. C. Schudlo and Chau 2014).
Word generation (WG)	Users think of as many words that start with a randomly generated letter. For example, given the letter "D", users may think of "dog, data, dashboard, donut…" (Ogata, Mukai, and Yagi 2007; Utsugi et al. 2007).
Happy thoughts (HT)	Users think of a past event in their life that made them happy (Kelly Tai and Chau 2009).
Relaxing with focus (RF)	Users concentrate on the trapezoid activation feedback. Focusing on either the increasing or decreasing portions (K. Izzetoglu et al. 2011).

Relaxing with slow counting (RS)	Users count slowly, starting from any number that they wish (Naseer and Hong 2013a).
Unconstrained rest (RR)	Users are allowed to let their mind wander and may think of anything other than the five mental tasks (Ang, Yu, and Guan 2012; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012; L. C. Schudlo and Chau 2014).

3.3.5 Additional data collection

The Ammons Quick Test was used to asses verbal IQ. The Ammons Quick Test is a 5-15 minute standardized verbal IQ test designed by Robert and Carol Ammons in 1962 and was administered after the last data collection session. The test consists of 50 questions in which users are asked to attribute a given word with one of four given pictures. The Ammons Quick Test has been shown to provide a good approximation of the full scale IQ as measured by the Wechsler intelligent scale for adults (WAIS) with Pearson's product moments of 0.85 (Zagar et al. 2013) and 0.89 (Husband and DeCato 1982). The Ammons Quick Test has been used in several psychiatric studies (Advokat, Eustis, and Pickering 2005; Marjoram et al. 2005). It should be noted that one of the ten participants (P4) chose not to complete the Ammons Quick Test.

A background questionnaire was administered prior to data collection to collect demographic data on each participant, including the participant's age range, gender, and handedness.

A post-session questionnaire was completed at the end of each session. Participants evaluated the following subjective statements on a 7-point Likert-type scale ranging from "Strongly Agree" to "Strongly Disagree": 1) I was tired before the session began, 2) I found it easy to concentrate during the session, 3) The session was fun, 4) It was easy to perform the tasks and session, and 5) The headgear was comfortable. For correlation analysis, the answers from all five post-session questionnaires were averaged for each participant.

3.3.6 Data processing

As a result of observed participant head motion or loss of contact between the head and the detectors, up to 20 data points (a maximum of four per class) were discarded from participants 3 and 9.

3.3.6.1 Filtering

The NIRS data were filtered in order to minimize noise due to the Mayer wave at a frequency of 0.1 Hz, the respiration cycle at a frequency of 0.2-0.4 Hz, and the cardiac cycle at a frequency of 0.5-2.0 Hz. A low-pass third-order Chebyshev infinite impulse response (IIR) cascade filter was used with a pass-band from 0-0.1 Hz, a transition band from 0.1-0.5 Hz, a stop-band from 0.5 Hz onwards, and a pass band ripple of 0.1 dB.

3.3.6.2 Calculating haemoglobin concentrations

The modified Beer Lambert's law (Delpy et al. 1988) was used to calculate changes in the concentrations of deoxygenated hemoglobin (Δ[Hb]), oxygenated hemoglobin (Δ[HbO]), and total hemoglobin (Δ[tHb]), as in (L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012; L. C. Schudlo and Chau 2014; Weyand, Takehara-Nishiuchi, and Chau 2015b; Weyand et al. 2015).

3.3.6.3 Feature extraction

Features extracted from the NIRS signal consisted of the temporal and spatial changes in the concentrations of the three chromophores (Hb, HbO, and tHb) over the four time windows (0-5s, 0-10s, 0-15s, and 0-20s). Specifically, the temporal features consisted of the linear regression slope over the normalized time windows for each of the 9 points of interrogation (4 time windows x 3 chromophores x 9 points of interrogation = 108 features), and the spatial features consisted of the linear regression slope over the zero to fourth order discrete orthogonal Chebyshev image moments over the four time windows (4 time windows x 3 chromophores x 15 image moments = 180 features). For more information on the extracted features, please refer to (L. C. Schudlo, Power, and Chau 2013; Weyand, Takehara-Nishiuchi, and Chau 2015c). Three distinct feature sets were considered during this study: temporal features only (108 features), spatial features only (180 features), and temporal combined with spatial (temporal-spatial) features (288 features).

3.3.6.4 Feature selection

For each of the three distinct feature sets, a subset of the features was selected from the training data to reduce the dimensionality of the problem and reduce redundancy. The Fast Correlation Based Filter (FCBF) was implemented (Yu and Liu 2003). FCBF is useful for feature sets with high dimensionality and has been used previously in EEG-BCI studies (Koelstra et al.

2010; Chanel, Ansari-Asl, and Pun 2007) and in the detection of the hemodynamic response by MRI (Tripoliti, Dimitrios, and Argyropoulou 2007). For more information on the FCBF please refer to (Yu and Liu 2003). The FCBF typically reduced the high dimensional feature sets to subsets consisting of 3 to 5 features.

3.3.7 Data analysis

3.3.7.1 Offline classification

Offline classification accuracies were calculated using ten iterations of ten-fold cross-validation (Refaeilzadeh, Tang, and Liu 2009). For each iteration of ten-fold cross-validation, the data were randomly separated into 10 equal sized portions (folds). Ten classification accuracies were calculated by iteratively using each fold as testing data and the remaining folds as training data. Only training data were used for feature selection and classifier training, and only the testing data were used to estimate the classification accuracies. Finally, all classification accuracies were averaged to estimate the overall accuracy.

3.3.7.2 Multi-class classification algorithm

Classification was performed for all possible n-class combinations of the six mental tasks (where n = 2, 3, 4 or 5 classes). Specifically, the 6 choose n (${}_6C_n$) task combinations for the 2-, 3-, 4-, and 5-class problems resulted in a total of 15, 20, 15, and 6 unique task combinations being explored for each classification problem, respectively.

Multi-class classification was conducted in a one-vs-one (OVO) manner by simplifying each n-class problem into m binary problems (where $m = {}_nC_2$) and voting on the majority class (Rocha and Goldenstein 2014; Dietterich and Bakiri 1995). The number of binary problems (m) given n = 2, 3, 4, and 5 classes, were 1, 3, 6, and 10, respectively.

The class of each of the m binary classifiers was determined by the majority vote of three ensemble classifiers, one for each feature set (temporal features only, spatial features only, and temporal and spatial features). In particular, a bagging ensemble classifier with ten members of linear discriminant analysis classifiers was used for each feature set.

Figure 11 shows an example of the classification algorithm for one set of 3 tasks (Task A vs. Task B vs. Task C). The label for the testing data was predicted by a majority vote of the three binary classifiers (A vs. B, A vs. C, and B vs. C), whose individual outputs were derived from a majority vote of three ensemble classifiers (temporal, spatial, and temporal-spatial).

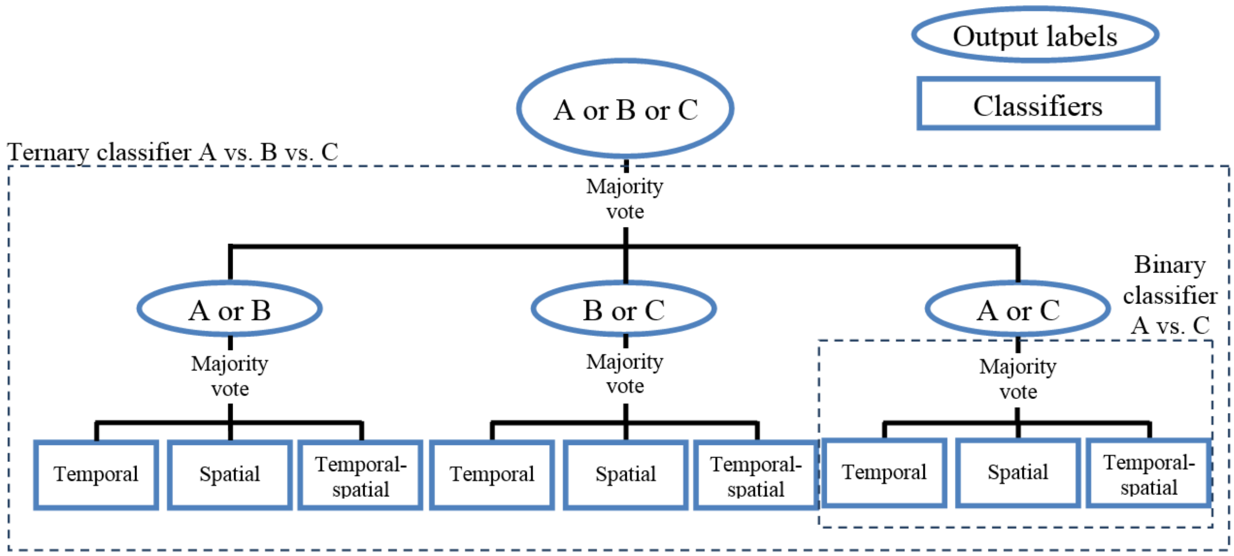

Figure 11. Classification scheme for a sample 3 class problem. The output of each of the binary (2-class) classifiers was determined by a majority vote of the three ensemble classifiers (temporal, spatial, and temporal-spatial). Subsequently, the output of the ternary (3-class) classifier was derived from the majority vote of the three binary classifiers.

3.3.7.3 Correlations between accuracy and user characteristics

The normality of the data was confirmed with the Shapiro-Wilk normality test. Pearson's correlations with $\alpha = 0.1$ were computed between accuracy and user characteristics following normal distributions, including IQ and state of mind data. Spearman's Rho correlations with $\alpha =$ 0.1 were computed between accuracy and the demographic data. For brevity, only the correlations with respect to the best 2-class accuracies were reported. The alpha value for the correlation analyses was set to 0.1 to minimize type II errors - i.e. missing a correlation that exists. Although this increases the probability of type I errors - i.e. finding a correlation when there isn't one - at this stage, we believe it is more important to find potential correlations (Banerjee et al. 2009). Moreover, a similar value has been used in a previous EEG-BCI correlation analysis (E V C Friedrich, Scherer, and Neuper 2013).

3.4 Results

3.4.1 Accuracies achieved

The accuracies achieved for the 2-, 3-, 4-, and 5-class problems are shown in Figure 12. Average classification accuracies of 78.4 ± 5.7%, 60.5 ± 6.6%, 46.7 ± 5.7%, and 37.2 ± 5.4%, were achieved for the 2-, 3-, 4-, and 5-class problems, respectively.

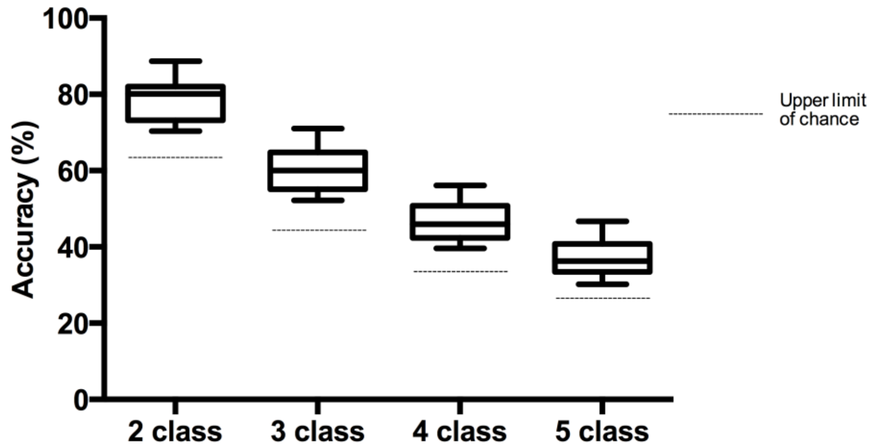

Figure 12. Box plot of accuracies for the 2-, 3-, 4-, and 5-class problems. Whiskers extend from min to max value. The dashed line below each box plot shows the upper limit of chance accuracy for each classification problem.

All participants exceeded the 70% threshold for the 2-class problem, and two participants (P3 and P5) exceeded the 70% threshold for the 3-class problem.

All participants exceeded chance levels for all *n*-class problems. Theoretically, for 2, 3, 4, and 5 classes, the chance level accuracies are 50%, 33%, 25%, and 20%, respectively. However, when the number of trials is less than infinity, the chance levels are those values plus or minus a confidence interval, given a value α (Mueller-Putz et al. 2008). Using the equation presented in (Mueller-Putz et al. 2008), the confidence limits for randomized class labels were calculated (Table 5). For the classifier accuracy to be statistically greater than chance, accuracies must exceed the upper confidence limit. In all cases, the classification of participant data with randomized class labels fell within the confidence limits of chance.

Table 5. Chance levels and corresponding confidence limits for the 2-, 3-, 4-, and 5-class problems given 60 trials per class and α = 0.05. The Bonferroni correction was used to account for the multiple task-subsets explored (Kaltenbach 2012)**. Therefore, adjusted α values of 0.0033, 0.0025, 0.0033, and 0.0083 were used for the 2-, 3-, 4-, and 5-class problems, respectively.**

Class	Chance Level	Confidence limits

2	50%	(36.8%, 63.2%)
3	33.4%	(23.2%, 44.2%)
4	25%	(17.2%, 33.6%)
5	20%	(14.3%, 26.5%)

3.4.2 Task frequency analysis

The best task pairs for each of the participants and for all n-class problems are shown in Table 6. The most common combination of tasks chosen for the 2-, 3-, 4-, and 5-class problems are shown in the last row of Table 6. Additionally, the individual task frequencies for each of the n-class problems are shown in Figure 13. The most frequently chosen tasks over all classification problems were happy thoughts (HT) and relaxing with focus (RF).

Table 6. Best task combinations for 2-, 3-, 4-, and 5-class problems. Legend: MM = mental math, WG = word generation, HT = happy thoughts, RF = relaxing with focus, RS = relaxing with slow counting, and RR = unconstrained rest.

Participant	2-class	3-class	4-class	5-class
1	WG RF	WG HT RF	MM WG HT RF	MM WG HT RF RS
2	MM RF	WG HT RF	MM WG RF RS	MM WG HT RF RS
3	HT RF	WG HT RR	WG HT RS RR	MM WG HT RF RR
4	MM WG	MM WG HT	MM WG HT RF	MM WG HT RF RS
5	MM RF	MM HT RF	MM HT RF RR	WG HT RF RS RR
6	HT RF	HT RF RR	MM HT RF RR	MM WG HT RF RR
7	WG RF	MM HT RF	MM WG HT RF	MM WG HT RF RS
8	MM HT	MM HT RF	MM HT RF RR	MM WG HT RF RS
9	HT RF	HT RF RR	MM HT RF RR	MM WG HT RF RR
10	MM RF	MM HT RF	MM HT RF RS	MM HT RF RS RR
Most common	HT&RF and MM&RF	MM&HT&RF	MM&HT&RF&RR	MM&WG&HT&RF&RS

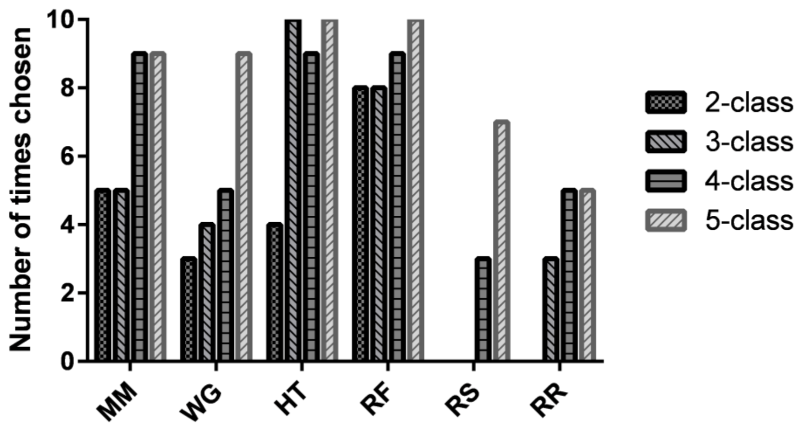

Figure 13. Number of times that each task was chosen as the best task for a participant for the 2-, 3-, 4-, and 5-class problems. It should be noted that since there are 10 participants, if a task is chosen 10 times, then it was chosen for all the participants. Legend: MM = mental math, WG = word generation, HT = happy thoughts, RF = relaxing with focus, RS = relaxing with slow counting, and RR = unconstrained rest.

3.4.3 Correlations of accuracy with user characteristics

Table 7 shows the correlations between assessed criteria and the 2-class classification accuracies over all participants. Strong positive correlations were found between accuracy and concentration (ρ = +0.634, $p < 0.05$), ease of session (ρ = +0.653, $p < 0.05$), and enjoyment (ρ = +0.550, $p < 0.10$), while accuracy and verbal IQ were strongly negatively correlated (ρ = -0.749, $p < 0.05$). It is noted that verbal IQ scores ranged from 80 to 116. Additionally, a moderate non-significant negative correlation was found between accuracy and tiredness before the session. Finally, weak to no correlation was found between accuracy and gender, as well as between accuracy and handedness. However, it is noted that there were 6 female participants and 8 right-handed participants. As a result of a homogeneous age range (seven of the participants were in their twenties), no correlation analysis between age and accuracy was conducted.

Table 7. Correlations between 2-class accuracies and user characteristics (α = 0.1).

	Correlation with 2-class accuracy (ρ)	**p-value (p)**
Verbal IQ	Pearson's ρ = -0.749	**0.020**
Enjoyment	Pearson's ρ = 0.550	**0.100**
Tiredness	Pearson's ρ = -0.449	0.193
Concentration	Pearson's ρ = 0.634	**0.049**
Ease of Session	Pearson's ρ = 0.653	**0.041**
Gender	Spearman's ρ = -0.221	0.540
Handedness	Spearman's ρ = 0.015	0.968

3.5 Discussion

3.5.1 Comparison of classification accuracies

The average 2-class accuracy achieved in this work (78.4%) appears to be on par with the high-end of those reported in other NIRS-BCI studies over the PFC (L. C. Schudlo, Power, and Chau 2013; S. Power, Kushki, and Chau 2011; Sarah D. Power, Falk, and Chau 2010). As expected, when moving beyond binary classification there was a significant drop in the accuracies achieved as a result of the increasing complexity of the classification problem. Overall, our results show promising progress towards distinguishing three and four mental tasks using an NIRS-BCI over the PFC. For most participants, the accuracies achieved in this study are still not sufficient for real-world BCI use; however, two participants (P3 and P5) were able to exceed the 70% threshold for a 3-class problem. It is noted that these two participants had the two highest enjoyment, concentration and reported ease-of-use ratings, and had two of the three lowest verbal IQ scores.

The average 3- and 4-class accuracies achieved in this work (61% and 47%) appear to be on par with the high-end of those reported in other multi-class NIRS-BCI studies over the PFC (Hirshfield et al. 2009; Sarah D Power, Kushki, and Chau 2012; Herff, Heger, Putze, et al. 2013; Herff, Heger, Fortmann, et al. 2013), namely, average 3-class accuracies of 50% (Herff, Heger, Fortmann, et al. 2013), 54% (Hirshfield et al. 2009), and 56% (Sarah D Power, Kushki, and Chau 2012), and a 4-class accuracy of 45% (Herff, Heger, Fortmann, et al. 2013). Moreover, contrary to our study, it appears that in NIRS-BCI literature, to date, no participant was able to exceed the 70% threshold for a 3-class problem. Additionally, to the best of our knowledge, no other NIRS-BCI study has attempted to differentiate five mental tasks over the PFC.

Note that the accuracies achieved for 3- and 4-class NIRS-BCIs over the PFC are still much lower than those achieved for EEG-BCIs and NIRS-BCIs over the motor cortex (Shin and Jeong 2014; An, Lee, and Ahn 2013; Gupta, Parameswaran, and Lee 2009; Palaniappan et al. 2002; Chai et al. 2012; Dobrea and Dobrea 2009). This is in line with similar trends of lower accuracies in 2-class NIRS-BCI studies over the PFC (S. Power, Kushki, and Chau 2011), when compared to 2-class EEG-BCI studies (Nai-Jen and Palaniappan 2004) and NIRS-BCI studies over the motor cortex (Sitaram et al. 2007). However, as previous researchers have pointed out, there are numerous advantages to using NIRS over the PFC. Specifically, the headband is not intrusive and requires minimal set-up time when compared to both EEG-BCIs and NIRS-BCIs over the motor cortex (Sarah D Power, Kushki, and Chau 2012; Bauernfeind et al. 2008; Kopton and Kenning 2014; Herff, Heger, Putze, et al. 2013). Additionally, motor tasks may not be suitable for all users, such as our target population of clients with motor impairments (E. Curran et al. 2004). When conducting NIRS-BCI measurements over areas with hair, there are additional challenges, including the integrity of the optode-skin contact and attenuation of light by hair (Lloyd-Fox, Blasi, and Elwell 2010). To combat this, spring loaded sources and detectors can be used, but these have been shown to be uncomfortable (Lloyd-Fox, Blasi, and Elwell 2010), with several participants dropping out of studies due to headset discomfort (Cui et al. 2011; Suzuki, Harashima, and Furuta 2010). On the other hand, it appears that users found the NIRS headband in this study to be comfortable. Participants evaluated the post-session statement “The headgear was comfortable” at an average rating of 5.3 +/- 0.9 on the 7 point Likert-type scale, indicating that on average participants agreed with this statement. Moreover, none of the participants reported the headset to be uncomfortable (no rating < 4).

3.5.2 Correlation between user characteristics and accuracy

Although correlation analyses were conducted on only ten participants, several interesting trends warrant further exploration.

3.5.2.1 Increasing accuracy with decreasing verbal IQ

Contrary to the EEG-BCI results by Hammer et al. (Hammer et al. 2012), who found no correlation between accuracy and non-verbal intelligence, in this study we found a strong negative correlation between accuracy and verbal intelligence. This correlation may seem surprising at first; however, upon further analysis, it appears that this trend may be attributable to task difficulty and neural efficiency.

In addition to the correlation between IQ and accuracy, a negative correlation was found between verbal IQ and the range (max-min) of HbO concentrations of the chosen tasks (ρ = -0.603, $p < 0.1$). This indicates that in general, individuals with lower IQ elicited larger overall changes in their hemodynamic activity. A possible explanation for this is that individuals with lower verbal IQ scores tend to elicit stronger, more consistent changes in neuronal hemodynamic activity when performing mental tasks since they find them to be more challenging.

In literature, the relationship between intelligence and hemodynamic brain activity is still widely debated, and is often referred to as the "neural efficiency debate" (Graham et al. 2010). Similar to findings in this study, several researchers reported that that a decrease in IQ or skill was associated with an increase in brain activation. An MRI study conducted by Graham et al. found that participants with average IQ showed greater PFC activation during response selection than did high IQ participants. The authors argued that the participants with high IQ's invoked more resource-efficient cognitive strategies resulting in less activation (Graham et al. 2010). Additionally, a positron emission tomography (PET) study by Haier et al., concluded that there was an inverse correlation between neural activity and verbal IQ scores in several brain regions, including areas of the frontal cortex (Haier et al. 1992). Moreover, an MRI study by Rypma et al. showed that during a working memory task, higher performing participants had overall less PFC activation than lower performing participants. This study also found that higher performing participants exhibited a larger increase in activation with increased task difficulty (Bart Rypma, Berger, and D'Esposito 2002; BART Rypma and D'esposito 1999). Collectively, these results reveal a relationship between task difficulty and IQ, and motivate further exploration of personalized task difficulty levels for each participant based on IQ scores.

3.5.2.2 Increasing accuracy with state of mind changes

We found a strong significant positive correlation between accuracy and each of concentration and reported ease. These results appear to be in line with EEG-BCI literature. Specifically, Hammer et al. noted a positive correlation between accuracy and attention ($\rho = 0.5$, $p = 0.02$) (Hammer et al. 2012), while Nijboer et al. documented a positive correlation between accuracy and mastery confidence ($b = 0.578$, $p < 0.05$) (Nijboer et al. 2008).

Additionally, we found a strong significant positive correlation between accuracy and task enjoyment. This finding also resonates with EEG-BCI literature. Friedrich et al. uncovered a positive correlation between accuracy and self-reported task enjoyment ($\rho = 0.3$, $p < 0.1$) (E V C Friedrich, Scherer, and Neuper 2013), while Nijboer et al. cited a positive correlation between accuracy and mood ($b = 0.498$, $p < 0.05$) (Nijboer et al. 2008).

Finally, we found a moderate, but not significant, negative correlation of accuracy with tiredness. This trend appears to be in contrast to the previous EEG-BCI finding of increased accuracy with decreased sleep in a P300 BCI ($p < 0.05$) (Guger et al. 2009). However, due to the very different neural mechanism involved in using a P300 BCI, this inconsistency is not surprising.

Overall, these findings motivate future research to enhance NIRS-BCI accuracy via training for confidence and concentration, and maximizing enjoyment while minimizing fatigue.

3.5.3 Limitations and future work

This study was conducted under controlled conditions which included a dimly lit room free of distractions. Future studies should be conducted in more practical environments in order to assess the functionality of the BCI in less than optimal conditions.

Secondly, the study was conducted on able-bodied participants. For use as an access technology for individuals with severe motor impairments, the results obtained likely do not reflect the performance of this target population. Further research and testing on a clinical population is necessary before conclusions about the effectiveness of multi-class BCIs can be made.

Thirdly, when using NIRS as an access modality for a BCI, there is the potential for systemic contributions to the signal (Takahashi et al. 2011). Although systemic noise is likely present, it has been shown that the majority of the signal originates from the cerebral cortex (Funane et al. 2014; Yoko Hoshi et al. 2011; Kirkpatrick et al. 1995). Moreover, the cortical component has been shown to be non-trivial; strong correlations have been reported between the NIRS and fMRI signals (Cui et al. 2011; Sato et al. 2013; Sasai et al. 2012) and between NIRS and EEG signals (Moosmann et al. 2003; Koch et al. 2008; Roche-Labarbe et al. 2010; Talukdar, Frost, and Diamond 2015).

Fourthly, in addition to the user characteristics described in this work, several other factors may be correlated to NIRS-BCI accuracy. Specifically, future work should explore the correlation of anatomical features with accuracy, such as scalp-cortex distance and frontal sinus volume, as these have been shown to be correlated with NIRS signal quality (Haeussinger et al. 2011). Other future directions include correlation analysis with respect to task performance, as well as the exploration of within-subject correlations on a per-session basis.

Finally, this study was conducted offline and over a relatively short period of time, with only five data collection sessions. It is possible that when moving online (with the inclusion of real-time performance feedback) and when conducting studies over longer periods of time (with the possibility of learning and habituation), performance and the correlation of performance with user characteristics may change. Further research on online and long-term trends is necessary.

3.6 Conclusions

This study explored the use of personalized mental tasks to increase the number of outputs of an NIRS-BCI. Average classification accuracies of 78.4 ± 5.7%, 60.5 ± 6.6%, 46.7 ± 5.7%, and 37.2 ± 5.4%, were attained for the 2-, 3-, 4-, and 5-class problems, respectively. All participants exceeded the 70% threshold for the 2-class problem, and most notably, two participants were able to exceed an accuracy of 70% for the 3-class problem.

Accuracy positively correlated with ease of session, ease of concentration, and enjoyment, and negatively correlated with verbal IQ. Future multi-class NIRS-BCI research ought to consider the development of training paradigms for maximizing user concentration, enjoyment, and confidence, as well as personalization of task difficulty based on IQ.

Overall, this research provides an incentive for further exploration of multi-class NIRS-BCIs, as well as continued research on the user characteristics that affect classification accuracies.

Chapter 4: User-selected Personalized Mental Tasks Online Study

The entirety of this chapter is reproduced from the article “Usability and performance-informed selection of personalized mental tasks for an online near-infrared spectroscopy brain-computer interface”. This manuscript has been published in Neurophotonics.

4.1 Abstract

Brain-computer interfaces (BCIs) allow individuals to use only cognitive activities to interact with their environment. The widespread use of BCIs is limited due in part to their lack of user-friendliness. The main goal of this work was to develop a more user-centered BCI and determine if: (i) individuals can acquire control of an online near-infrared spectroscopy BCI via usability and performance-informed selection of mental tasks without compromising classification accuracy, and (ii) the combination of usability and performance-informed selection of mental tasks yields subjective ease-of-use ratings that exceed those attainable with prescribed mental tasks. Twenty able-bodied participants were recruited. Half of the participants served as a control group, using the state-of-the-art prescribed mental strategies. The other half of the participants comprised the study group, choosing their own personalized mental strategies out of eleven possible tasks. It was concluded that users were in fact able to acquire control of the more user-centered BCI without a significant change in accuracy compared to the prescribed task BCI. Furthermore, the personalized BCI yielded higher subjective ease-of-use ratings than the prescribed BCI. Average online accuracies of 77 ±12.9% and 73 ±12.9% were achieved by the personalized and prescribed mental task groups, respectively.

4.2 Introduction

4.2.1 Brain-computer interfaces

Brain-computer interfaces (BCIs) allow individuals to interact with their environment using only cognitive activities (Elisabeth V C Friedrich, Scherer, and Neuper 2012; Ang, Yu, and Guan 2012; S. M. Coyle, Ward, and Markham 2007; Niels Birbaumer 2006). BCIs exploit a user’s brain signals for external device control without requiring intentional muscle activations or the peripheral nervous system responses (S. M. Coyle, Ward, and Markham 2007; J. Wolpaw et al. 2000). The potential applications of BCIs are vast. BCIs can be used by able-bodied individuals for gaming, entertainment, and to accelerate learning (Elisabeth V C Friedrich, Scherer, and Neuper 2012; Ayaz et al. 2007). BCIs can also be used by individuals with severe motor impairments as a means of communication, as a way of controlling a wheel chair for mobility, or as a method for controlling devices in their environment (Niels Birbaumer 2006; J. Wolpaw et al. 2000; Ayaz et al. 2007; Blain, Mihailidis, and Chau 2008). Individuals with amyotrophic lateral sclerosis, spinal cord injuries, brain stem stroke, complete paralysis, or muscular dystrophy among other debilitating conditions stand to benefit from BCIs (Ayaz et al. 2007; Sitaram et al. 2007). Indeed, BCIs have the potential to significantly increase the quality of life for patients with severe motor impairments (Naito et al. 2007).

A BCI consists of an input, a translation algorithm, and an output. The input to a BCI is the physiological signal that is being harnessed. The input can be further broken down into the access modality, which refers to how the signal is collected, and the access pathway, which refers to how a change in that signal is evoked. After the signal is collected, the translation algorithm processes the signal to remove noise, extracts features, extracts the most discriminant features, and trains a classifier to predict the class to which a case belongs. Finally, the output of the BCI involves the control of an external device (J. Wolpaw et al. 2000). For a given BCI, any of the above factors can be modulated in order to improve the BCI. Most papers focus on improving the translation methods, while only a few papers focus on improving the input. The focus of this research is on improving the input access pathway.

4.2.2 Near-infrared spectroscopy

Near-infrared spectroscopy (NIRS) is a non-invasive, safe, portable, and low-cost optical neural imaging technique that measures haemodynamic brain activity (Ang, Yu, and Guan 2012; S. M. Coyle, Ward, and Markham 2007; Ayaz et al. 2007; S. Coyle et al. 2004). Despite temporal limitations, NIRS offers several advantages over other non-invasive BCI access modalities, including for example, gel and paste-free donning and flexibility of measurement environments. For further discussion of the relative merits of NIRS as an access modality, refer to Refs. (S. M. Coyle, Ward, and Markham 2007; Ayaz et al. 2007; Sitaram et al. 2007; S. Coyle et al. 2004). NIRS works by measuring the changes in the absorption of near infrared light that travels through the skin, periosteum, skull, meninges, and the cerebral cortex of the brain. The amount of light that is absorbed varies with the amount of oxygen in the blood. Through a mechanism known as neurovascular coupling, areas of the brain that are active have an increase in oxygenated haemoglobin (HbO), an increase in total haemoglobin (tHb), and a decrease in deoxygenated haemoglobin (Hb) (Wolf et al. 2002; Arno Villringer and Chance 1997; Meltem Izzetoglu et al. 2007). However, other coupling trends have also been reported (Bauernfeind et al. 2008; A Villringer et al. 1993; Gert Pfurtscheller, Bauernfeind, et al. 2010; Quaresima et al. 2005; Y Hoshi et al. 1994; Koshino et al. 2011; Buckner, Andrews-Hanna, and Schacter 2008). NIRS provides a indirect measure of cognitive activity by ascertaining the changes in the concentration of HbO and Hb in the brain (Ang, Yu, and Guan 2012; S. M. Coyle, Ward, and Markham 2007).

4.2.3 Prescribed mental tasks

NIRS is a promising access modality; however, to date, little research has been done on the access pathways accompanying this access modality (Dobrea and Dobrea 2009). Currently, to the best of our knowledge, most NIRS-BCI studies have used prescribed mental activation tasks. The tasks used to control the BCI are chosen by researchers based on previous studies showing differentiability in the activation or deactivation caused by a specific set of tasks. By discriminating between the changes in the NIRS signal resulting from the user performing each task, one is able to control the binary BCI. Several different mental tasks have been used in past NIRS-BCI studies, including: mental math (Ang, Yu, and Guan 2012; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Ogata, Mukai, and Yagi 2007; Bauernfeind et al. 2008; Sarah D. Power, Kushki, and Chau 2012; Utsugi et al. 2007; L. C. Schudlo and Chau 2014), mental singing (Naito et al. 2007; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011), word generation (Herff, Heger, Putze, et al. 2013; Ogata, Mukai, and Yagi 2007; Utsugi et al. 2007), memory (Ayaz et al. 2007; Ogata, Mukai, and Yagi 2007; Utsugi et al. 2007), mental counting (Naseer and Hong 2013a), mental rotation (Herff, Heger, Putze, et al. 2013), concentration (K. Izzetoglu et al. 2011), motor imagery (S. M. Coyle, Ward, and Markham 2007; Sitaram et al. 2007; S. Coyle et al. 2004; Kanoh et al. 2009; Naseer and Hong 2013b), and rest (Ang, Yu, and Guan 2012; Ayaz et al. 2007; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012; Naseer and Hong 2013a; L. C. Schudlo and Chau 2014).

4.2.4 Motivation for user-selected personalized mental tasks

An alternative to using prescribed mental tasks is to use personalized mental tasks, where each user has a set of tasks selected specifically for him or her. To date, the exploration of personalized mental tasks in NIRS-BCIs is limited. However, personalized tasks have been explored in MRI (Sorger et al. 2009) and EEG (Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012) BCI studies. It appears that in all MRI-BCI and EEG-BCI personalized task studies, the actual tasks were selected by the researcher based solely on performance with the aim of improving the BCI accuracy. Overall, these studies conclude that there is significant inter-subject variability in brain activation elicited by the same mental tasks and cognitive processes, and as a result the tasks that are most effective for controlling a BCI vary among users (Sorger et al. 2009; Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012).

Although improvements in accuracy are important, improving the BCIs usability has also been concluded in literature to be vitally important. In a review of the first international meeting devoted to BCI research and development, Wolpaw et al. described that the widespread application of BCI-based communication systems will depend on cost, ease of training, ease-of-use, and user satisfaction (J. Wolpaw et al. 2000). Additionally, Bos et al. found that ease-of-use was the second most important factor after accuracy in BCI acceptance (Bos, Poel, and Nijholt 2011; Tan and Nijholt 2010). Furthermore, a study by Holz et al. found that ease-of-use was one of the most important aspects of the BCI for four severely motor-restricted end-users (Holz et al. 2013). From these studies, it can be concluded that BCI usability is greatly important and even if classification accuracies are very high, if users dislike performing their assigned tasks, they are not likely to use the BCI (J. Wolpaw et al. 2000; Bos, Poel, and Nijholt 2011). Improving the ease-of-use of a BCI could result in increased user satisfaction and user-friendliness, which could lead to increased adoption and decreased BCI abandonment (J. Wolpaw et al. 2000; Bos, Poel, and Nijholt 2011; Tan and Nijholt 2010).

Currently, when using prescribed tasks or personalized tasks chosen solely based on accuracy, users often find the assigned tasks not suitable, unenjoyable, or difficult to perform (Bos, Poel, and Nijholt 2011; E. Curran et al. 2004; E V C Friedrich, Scherer, and Neuper 2013). For example, math tasks may not be suitable for individuals who have difficulty with, minimal knowledge of, or a dislike for arithmetic (E. Curran et al. 2004). Personalized mental tasks that are chosen by the user based on both performance and usability could result in the development of a more user-centered BCI that is easier to use and more enjoyable.

We previously conducted an offline study to compare four mental task frameworks: two user-selected personalized mental task frameworks, a researcher-selected personalized mental task framework, and a prescribed mental task framework (Weyand, Takehara-Nishiuchi, and Chau 2015c). We showed that user-selected personalized tasks have the potential to yield higher perceived ease-of-use ratings (Weyand, Takehara-Nishiuchi, and Chau 2015c). However, further studies are needed to verify the value of personalized tasks by comparing personalized and prescribed task selection schemes, particularly in an online paradigm. To the best of our knowledge, no other NIRS-BCI study has explored the use of personalized tasks and no other BCI studies have allowed users to choose personalized tasks based on both performance and ease-of-use.

4.2.5 Objectives

The aim of this research was to develop a more user-centered personalized mental task access pathway for an NIRS-BCI that allows individuals to choose tasks based on their performance and subjective ease-of-use ratings. This study pursues two research questions: (1) Can individuals acquire control of an online NIRS-BCI via usability and performance-informed selection of mental tasks without compromising classification accuracy? (2) Can the combination of usability and performance-informed selection of mental tasks for NIRS-BCI control yield subjective ease-of-use ratings that exceed those attainable with prescribed mental tasks?

4.3 Methods

4.3.1 Participants

Twenty able-bodied subjects (eight male) between the ages of 16 and 40 were recruited from the staff and students at Holland Bloorview Kids Rehabilitation Hospital (Toronto, Canada). All participants were right-handed according to the Edinburgh handedness test (Oldfield 1971). Participants had normal or corrected-to-normal vision and had no known motor impairments, degenerative disorders, cardiovascular disorders, trauma-induced brain injuries, drug or alcohol-related conditions, psychiatric conditions, respiratory disorders or metabolic disorders. Participants were asked to not smoke or drink alcoholic or caffeinated beverages three hours prior to each data collection session. The study was conducted with informed consent and with ethics approval from the Holland Bloorview Kids Rehabilitation Hospital and the University of Toronto.

Half of the participants were randomly allocated to the study group, which chose their own personalized mental tasks, and the other half to the control group, which were assigned prescribed mental tasks. Since gender (Marumo et al. 2009; Yang et al. 2007; F. Okada et al. 1993; Tamura, Hoshi, and Okada 1997), handedness (F. Okada et al. 1993; Tamura, Hoshi, and Okada 1997), and age (Schroeter et al. 2003; Kwee and Nakada 2003; Arlene Duncan et al. 1996) have been shown to affect NIRS measurements, the two groups were matched based on these three criteria. One participant (male) from the personalized mental task group was excluded from all results analysis since he was not able to follow the experimental protocol. It is noted that the participants in the study group went on to partake in ten more sessions. Data from those sessions were not used in the present analysis; however, it was used to explore self-regulation as an alternative NIRS-BCI access pathway in Ref. (Weyand, Takehara-Nishiuchi, and Chau 2015b).

4.3.2 Experimental setup

The NIRS data were collected using a multi-channel frequency-domain NIRS system with a sampling rate of 31.25 Hz (Imagent Functional Brain Imaging System from ISS Inc., Champaign, IL (ISS Inc. 2012)). The NIRS system was used to measure the blood oxygen content from the prefrontal cortex (PFC). The PFC is involved in higher brain functions, including logical thinking, planning, and emotion (Ogata, Mukai, and Yagi 2007; Gao et al. 1990).

Five laser diodes each emitting 690 nm and 830 nm light and three photomultiplier tube detectors attached to a headband were used. The headband was made out of a rubber polymer (3M 9900 series), which is comfortable on the skin and easily conforms to the shape of any head. Black fabric was sewn on the outside of the headband, to create tight, opaque pockets within which the light sources and detectors were fit. These pockets secured the sources and detectors, minimizing their motion while maximizing their contact with the head. The headband was centered on the participant’s forehead with reference to the nose, and was placed directly above the eyebrows, as shown in Figure 14 (A).

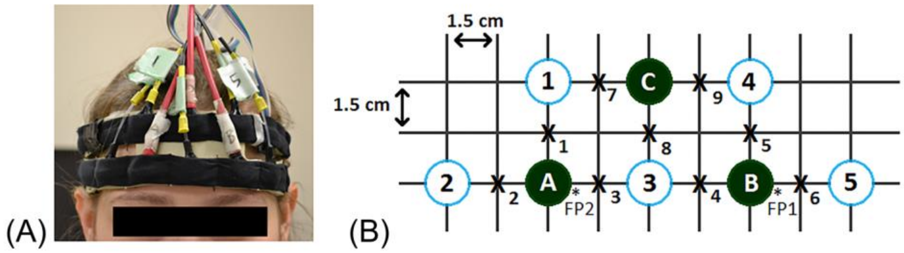

Figure 14. (A) NIRS headband placed over the forehead. (B) Experimental source and detector configuration. Legend: the solid circles represent detectors; the open circles represent light source pairs; the x's represent points of interrogation; and the starred areas represent the FP1 and FP2 positions of the international 10-20 EEG system.

The sources and detectors were arranged in a trapezoidal shape. Each source and adjacent detector was separated by a distance of 3 cm. This distance corresponds to a penetration depth of approximately 2.5 cm, which has been shown to reach the outer layer of the cerebral cortex (Bauernfeind et al. 2008; Haeussinger et al. 2011; E. Okada et al. 1997). Several other NIRS-BCI studies have also used a source-detector separation distance of 3 cm over the PFC (Naito et al. 2007; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Bauernfeind et al. 2008; Sarah D. Power, Kushki, and Chau 2012). The source-detector configuration allowed for the interrogation of nine discrete locations. A schematic diagram of the configuration and points of interrogation are shown in Figure 14 (B).

Neurofeedback was provided during all sessions in the form of: 1) a trapezoid topographic image showing the real-time changes in blood oxygenation levels over the PFC and 2) a ball that rose and fell with the average change over the entire interrogation area. The feedback was updated every 125 ms, and was calculated using cubic interpolation of the HbO values between the points of interrogation. HbO was chosen for the feedback since it has been cited to be more indicative of activity than Hb and tHb (S. M. Coyle, Ward, and Markham 2007; S. Coyle et al. 2004). Participants were informed that the red colour on the feedback represented an increase in haemodynamic activity, while the blue colour represented a decrease in haemodynamic activity. The activation feedback is shown in Figure 15.

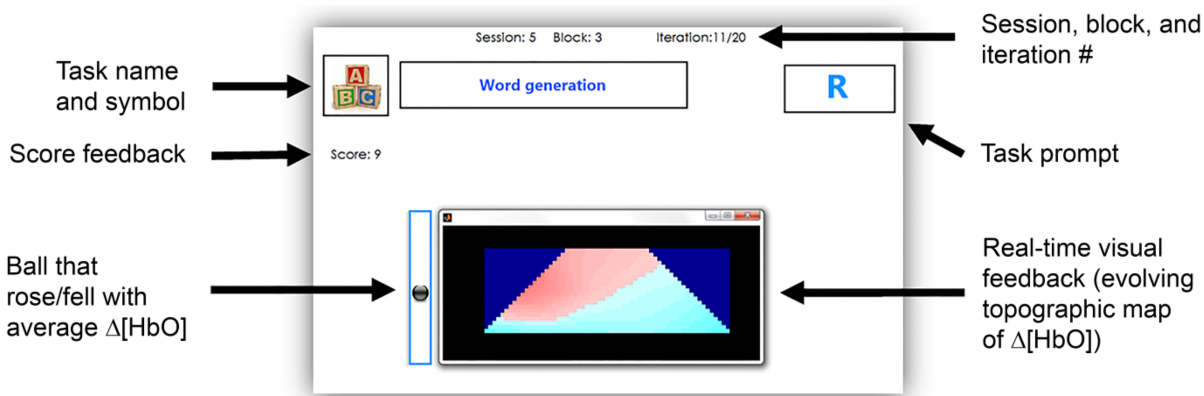

Figure 15. User Interface for Session 5, Blocks 2 and 3 (online classification with score feedback).

4.3.3 Personalized task measures

In this study, 11 possible mental tasks were considered based on their deployment in previous BCI studies or their documented ability to induce PFC activity in functional imaging studies. Each of the eleven tasks is described in Table 8.

Table 8. Eleven mental tasks used in sessions 1 to 3.

Task	Description
Mental math (Math)	Participants were prompted with a math problem that appeared in the top right corner of the screen, and they were asked to repeatedly subtract a two digit number from a three digit number. For example, given 986-12; the participant would mentally evaluate 986-12 = 974; 974-12 = 962; 962-12 = 950; and so on. Numbers were randomly generated. This task has been used in several previous NIRS-BCI studies (Ang, Yu, and Guan 2012; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Ogata, Mukai, and Yagi 2007; Bauernfeind et al. 2008; Sarah D. Power, Kushki, and Chau 2012; Utsugi et al. 2007; L. C. Schudlo and Chau 2014; Naseer, Hong, and Hong 2014; Khan, Hong, and Hong 2014).
Mental singing (Music)	Participants were asked to sing a song in their head. They were informed they could choose to sing any song they wanted, but they were asked to pick a song that they liked. This task has been used in previous NIRS-BCI studies (Naito et al. 2007; Herff, Heger, Putze, et al. 2013; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011).
Word generation (Words)	Participants were asked to think of as many words as possible that start with a specific letter. For example, if the letter "D" appeared on the screen, the user may think of the words: dog, draft, door, deli, and so on. Letters (excluding x and z) were randomly generated. This task has been used in previous NIRS-BCI studies (Ogata, Mukai, and Yagi 2007; Utsugi et al. 2007; Faress and Chau 2013).
Tangram puzzle (Rotation)	Participants were prompted with a tangram puzzle in the top right corner of the screen, and were asked to imagine rotating the pieces to make a final picture. This task was chosen because it has been shown to alter PFC activity, measured using NIRS in a non-BCI study (Ayaz et al. 2012). A similar task was also used in a previous NIRS-BCI study (Ayaz et al. 2012; Herff, Heger, Putze, et al. 2013).
Relaxing with	Participants were asked to slowly count in their heads while relaxing. A similar task

counting (Counting)	has been used in a previous NIRS-BCI study (Naseer and Hong 2013a; Khan, Hong, and Hong 2014).
Happy thoughts (Happy)	Participants were asked to think about the details of a past event in their life that made them very happy. This task has been used in a previous NIRS-BCI study: (Kelly Tai and Chau 2009). This task also uses episodic memory, which has been shown to alter activity in the PFC (Koshino et al. 2011).
Word colour (Stroop)	Participants were prompted with a series of colour names. The words were also coloured, but the colour of the words did not always match the written word. For example, the word blue may have been written in red ink. The participants were asked to say the real colour of the word in their head. This task is commonly referred to as the stroop task. This task has been used in a previous NIRS non-BCI study (Schroeter et al. 2002; Ehlis et al. 2005).
Visualizing the future (Future)	Participants were asked to imagine their life in five years, specifically focusing on their future day-to-day activities. This task was chosen for its potential to alter activity in the PFC. The PFC is part of the default mode network and has been shown to be activated when envisioning the future and during self-relevant mentalization (Buckner, Andrews-Hanna, and Schacter 2008).
Relaxing with focus on the feedback (Focus)	Participants were asked to relax and focus on the feedback. A similar task has been used in a previous NIRS-BCI study (K. Izzetoglu et al. 2011).
Motor imagery (Motor)	Participants were asked to imagine moving their arms or legs. Motor imagery has been investigated in previous NIRS-BCI studies, but strictly over the motor cortex (Sitaram et al. 2007; S. M. Coyle, Ward, and Markham 2007; Naseer and Hong 2013b). This task was included since it has been shown that the PFC is also affected by motor imagery (Kanthack, Bigliassi, and Altimari 2013; Leff et al. 2011).
Rest (Rest)	Participants were asked to relax and let their minds rest. This task has been used in previous NIRS-BCI studies (Ang, Yu, and Guan 2012; Ayaz et al. 2007; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012; Naseer and Hong 2013a; L. C. Schudlo and Chau 2014; Naseer, Hong, and Hong 2014).

To facilitate a user-centered approach to task selection that allows one to strike a personal balance between usability and performance, two task measures were invoked, namely, a total ease-of-use score and a weighted slope score.

Total ease-of-use score: Subsequent to performing an iteration of each task, users rated their perceived ease-of-use on a 5-point Likert-type scale, ranging from “very easy” to “very difficult”, as per recommendations for measuring post-task usability (Tedesco and Tullis 2006; Sauro and Dumas 2009). A total ease-of-use score for each task was determined as the average ease-of-use rating across all iterations of the task.

Weighted slope score: Task performance was captured by a task-specific weighted slope score (WS-Score) that represents the ability of a subject to consistently increase or decrease their haemodynamic activity by performing a task. Specifically, the weighted slope score, WS_i, for the i^{th} task, was defined as

$$WS_i = \frac{\frac{1}{N}\sum_{j=1}^{N} m_{ij}}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(m_{ij} - \left[\frac{1}{N}\sum_{k=1}^{N} m_{ik}\right]\right)^2}} \quad (2)$$

where m_{ij} and m_{ik}, are the slopes of the least square line of best-fit to the haemodynamic activity over time for the j^{th} or k^{th} iteration of the i^{th} task, and N is the number of times the task was performed. For each iteration j or k, of each task, i, an average response is computed from the haemodynamic response ($\Delta[HbO]$) from each of the 9 channels (i.e., 9 points of interrogation). A slope value, m_{ij} or m_{ik}, is extracted from the best-fit line to each channel-averaged response. The weighted slope score is then computed as the mean of all slopes for each iteration of a given task divided by the corresponding standard deviation, providing a measure of the tendency for a task to consistently increase or decrease haemodynamic activity.

4.3.4 Data collection sessions

All participants took part in five data collection sessions on five separate days, spanning a period of one week. In each session, participants were seated in front of a computer in a dimly lit room. The general protocol was the same for all sessions. Each session started with a short warm-up period, which allowed the user to become familiar with the interface. Following the warm-up, each participant took part in three data collection blocks. During each data collection block, the participant performed either 22 task intervals (session 1, 2, and 3) or 20 task intervals (session 4 and 5). Each task interval involved a task being performed for 17 seconds, followed by a 20-second rest. The duration of the task and rest activities were chosen on the basis of preliminary data and past work (S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013). A schematic illustration of the overall study, session, and block structure is shown in Figure 16.

Study Structure

Sessions 1-3 Perform all eleven tasks			Session 4 Control group: Prescribed tasks Study group: Personalized tasks			Session 5 Control group: Prescribed tasks Study group: Personalized tasks		
Block 1	Block 2	Block 3	Block 1	Block 2	Block 3	Block 1	Block 2	Block 3
Offline			Offline			Offline	Online	Online

Block Structure

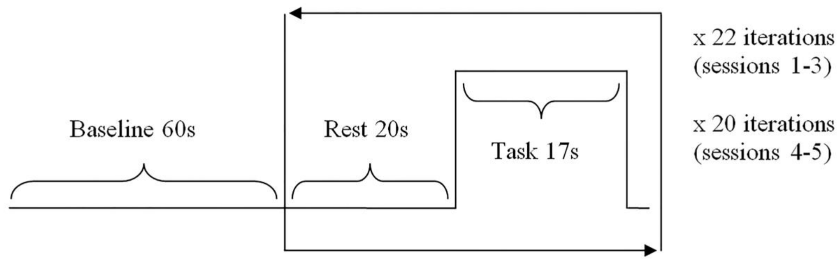

Figure 16. Study, session, and block structure.

Sessions 1-3: Participants performed each of the eleven tasks twice per block. The tasks were presented in a random order. Immediately after performing each task and before the twenty second rest, the user was asked to rate the task in terms of its ease-of-use and desirability for BCI control. The twenty second rest period allowed cortical haemodynamic changes from the previous task and the ease-of-use scoring to dissipate. The goal of these three sessions was to determine the magnitude of change in blood oxygenation when the participants performed each task and to determine the level of subjective enjoyment of each task. By the end of the third session, participants had performed each task 18 times (3 sessions x 3 blocks x 2 iterations), and thus $N = 18$, in Equation (2) above.

Session 4: The control group was assigned the mental math and rest tasks irrespective of their performance in the first three sessions. These tasks represent the current state-of-the-art in NIRS-BCIs (Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012). On the other hand, the study group was instructed to choose their own pair of personalized tasks from among the 11 possibilities. To inform their choice, participants were provided with their total ease-of-use rating for each task from sessions 1-3. In addition, for each task, participants were presented with line graphs depicting HbO concentration changes averaged over the 9 points of interrogation. Graphs for the top three tasks for increasing haemodynamic activity (highest WS-Scores) and the best three tasks for decreasing the haemodynamic activity (lowest WS-Scores) were shown. Participants were asked to choose a preferred task that tended to increase their haemodynamic response - their ‘increasing task’ - and another that tended to decrease their haemodynamic response - their ‘decreasing task’. Subsequently, participants were prompted to perform their two tasks: mental math and rest for the comparison group, and personalized tasks for the study group.

Session 5: In the first block, participants were prompted to perform their two tasks as in Session 4. During the second and third blocks of the fifth session, participants continued to perform their two tasks when prompted by the interface (Figure 15) but received the corresponding label (i.e. increase or decrease task) estimated by the computer. A score counter was updated, displaying the number of times that the computer correctly labeled (classified) the task performed by the user. On average classification took one to two seconds; however, this will vary depending on the speed of the computer. An example of the user interface from session five, block three, is shown in Figure 15.

4.3.5 Signal treatment

4.3.5.1 Filtering

NIRS data are affected by various sources of physiological noise. In particular, the NIRS signal is contaminated with the Mayer wave at a frequency of 0.1 Hz, the respiration cycle at a frequency of 0.2 Hz to 0.4 Hz, and the cardiac cycle at a frequency of 0.5 Hz to 2 Hz (Sarah D. Power, Falk, and Chau 2010; Ayaz et al. 2009). To mitigate the influences of these noise sources, the NIRS signal was passed through a digital low-pass filter in real-time using a third-order Chebyshev infinite impulse response (IIR) cascade filter with a pass-band from 0 to 0.1 Hz, a transition band from 0.1 to 0.5 Hz, a stop-band from 0.5 Hz onwards, and a pass band ripple of 0.1 dB.

4.3.5.2 Calculating haemoglobin concentrations

After filtering the data, the changes in concentrations of HbO, Hb, and tHb, i.e $\Delta[HbO]$, $\Delta[Hb]$, and $\Delta[tHb]$, were calculated using the modified Beer-Lambert’s Law (S. M. Coyle, Ward, and Markham 2007; J. Wolpaw et al. 2000; Niels Birbaumer and Cohen 2007; Delpy et al. 1988).

$$\Delta[HbO] = \frac{\varepsilon_{Hb}^{\lambda_2}\left(\frac{\log\left(\frac{I_B^{\lambda_1}}{I_A^{\lambda_1}}\right)}{DPF^{\lambda_1}}\right) - \varepsilon_{Hb}^{\lambda_1}\left(\frac{\log\left(\frac{I_B^{\lambda_2}}{I_A^{\lambda_2}}\right)}{DPF^{\lambda_2}}\right)}{r\left(\varepsilon_{Hb}^{\lambda_2}\varepsilon_{HbO}^{\lambda_1} - \varepsilon_{Hb}^{\lambda_1}\varepsilon_{HbO}^{\lambda_2}\right)} \quad (3)$$

$$\Delta[Hb] = \frac{\varepsilon_{HbO}^{\lambda_2}\left(\frac{\log\left(\frac{I_B^{\lambda_1}}{I_A^{\lambda_1}}\right)}{DPF^{\lambda_1}}\right) - \varepsilon_{HbO}^{\lambda_1}\left(\frac{\log\left(\frac{I_B^{\lambda_2}}{I_A^{\lambda_2}}\right)}{DPF^{\lambda_2}}\right)}{r\left(\varepsilon_{Hb}^{\lambda_1}\varepsilon_{HbO}^{\lambda_2} - \varepsilon_{Hb}^{\lambda_2}\varepsilon_{HbO}^{\lambda_1}\right)} \quad (4)$$

$$\Delta[tHb] = \Delta[Hb] + \Delta[HbO] \quad (5)$$

where I_B^{λ} is the mean light intensity measured at baseline at wavelength λ, I_A^{λ} is the light intensity measured at any given time at wavelength λ, $\varepsilon_{Hb}^{\lambda}$ and $\varepsilon_{HbO}^{\lambda}$ are the specific extinction coefficient of deoxygenated and oxygenated haemoglobin respectively, at wavelength λ, DPF^{λ} is the differential path factor at wavelength λ, and r is the geometric distance between the emitter and detector. For a derivation of these equations see Refs. (Delpy et al. 1988; Kelly Tai and Chau 2009). In this study, r = 3cm, $\varepsilon_{690nm,Hb}$ = 2.1382 $mM^{-1}cm^{-1}$ (Cope 1991), $\varepsilon_{830nm,Hb}$ = 0.7804 mM^{-1} cm^{-1} (Cope 1991), $\varepsilon_{690nm,HbO}$ = 0.3123 mM^{-1} cm^{-1} (Cope 1991), $\varepsilon_{830nm,HbO}$ = 1.0507 mM^{-1} cm^{-1} (Cope 1991), DPF_{690nm} = 6.51 (A Duncan et al. 1995), and DPF_{830nm} = 5.86 (A Duncan et al. 1995).

4.3.5.3 Feature extraction and feature selection

Features were extracted over four time windows (0-5s, 0-10s, 0-15s, and 0-17s). Features included the temporal changes in the three chromophores (Hb, HbO, and tHb) at each of the 9 points of interrogation (4 time windows x 3 chromophores x 9 points of interrogation = 108 features) and the spatiotemporal changes of the zero to fourth order discrete orthogonal Chebyshev image moments (4 time windows x 3 chromophores x 15 image moments = 180 features) (L. C. Schudlo, Power, and Chau 2013). A total of 288 features were thus extracted from the data.

The temporal feature extraction involved normalizing each measured response, by subtracting the mean and dividing by the standard deviation, and then determining the least square line of best-fit slope of the concentration change of each of the three chromophores over each of the four time windows. For example, the first feature was the least square line of best-fit slope of [Hb] at the first point of interrogation, over the first five seconds that the task was performed. Temporal features have been previously deployed with NIRS-BCIs (Sarah D Power, Kushki, and Chau 2012; Sarah D. Power, Kushki, and Chau 2012; L. C. Schudlo, Power, and Chau 2013; L. C. Schudlo and Chau 2014; Sarah Dianne Power and Chau 2013).

To derive the spatial features, topographic images for $\Delta[HbO]$, $\Delta[Hb]$, and $\Delta[tHb]$ were generated by spatial interpolation of the data at the nine points of interrogation. Specifically, cubic interpolation at equally spaced intervals between the points of interrogation was performed, as in Ref. (L. C. Schudlo, Power, and Chau 2013), to create a trapezoidal image, 21 pixels in height and with parallel sides 21 and 61 pixels in length. To account for inter-trial variability, the pixel values were normalized to fall between 0 and 1, as in Ref. (L. C. Schudlo, Power, and Chau 2013). To summarize the spatial changes, zero to fourth order discrete orthogonal Chebyshev polynomial image moments were extracted from each image at every instance in time. Image moments are a weighted average of the image pixel intensities and take the general form:

$$M_{mn} = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} P_m(x) P_n(y) f(x,y) \quad (6)$$

where, $f(x,y)$ is the intensity distribution of the N_x by N_y image, x and y are the pixel coordinates, m and n are the degrees (orders) of the Chebyshev polynomials, $m+n$ is the moment order, and $P_m(x)$ and $P_n(y)$ are the two-dimensional orthogonal Chebyshev polynomials (Zhu et al. 2010), calculated using equation 9, equation 12, and Table 2 from Ref. (Zhu et al. 2010). A total of 15 image moments were extracted from each image at each time point, one for each possible permutation of moment orders of 0, 1, 2, 3, and 4, as shown in Table 9. Finally, the simple least square line of best-fit slope of each image moment signal was calculated over each time window. For example, the first feature was the least square line of best-fit slope of the change of the zeroth order moment $(m = n = 0)$ of [Hb] over the first five seconds that the task was performed.

Table 9. Degree *m* and *n* of each of the fifteen image moments.

Image moment number	Moment order (*m*+*n*)	*m*	*n*
1	0	0	0
2	1	0	1
3	1	1	0
4	2	0	2
5	2	2	0
6	2	1	1
7	3	0	3
8	3	3	0
9	3	1	2
10	3	2	1
11	4	0	4
12	4	4	0
13	4	1	3
14	4	3	1
15	4	2	2

The Sequential Forward Floating Search (SFFS) algorithm was used to select a subset of eight features from the total feature set used for classifier training (L. C. Schudlo, Power, and Chau 2013; Pudil, Novovičová, and Kittler 1994). In general, SFFS uses a criterion function to assess the discriminative capabilities of each candidate feature set. Starting with an empty feature set, the algorithm sequentially adds features with the largest criterion value. At each iteration, the method also removes a previously added feature that is presently the least significant with respect to the criterion function (Pudil, Novovičová, and Kittler 1994). We used the Fisher criterion to assess the discriminatory capabilities of each feature set, as in Refs. (L. C. Schudlo and Chau 2014; L. C. Schudlo, Power, and Chau 2013). For this study, the target number of eight features was chosen on the basis of preliminary data and past work (Sarah D Power, Kushki, and Chau 2012; L. C. Schudlo and Chau 2014).

4.3.5.4 Pattern classification

An ensemble of three classifiers was used to differentiate between task-induced changes in the haemodynamic response, as in Ref. (L. C. Schudlo, Power, and Chau 2013). In particular, for each participant, three linear discriminant analysis classifiers (LDAs) were trained. In general, LDAs seek to separate classes by projecting the training samples onto a line that maximizes class separabilty (Bishop 2006). The first classifier was trained using eight features selected from the 108 temporal features; the second classifier was trained using eight features selected from the 180 spatiotemporal features; and the third classifier was trained using eight features selected from all 288 features (temporal and spatiotemporal). Feature extraction and selection are described in Sec. 4.3.5.3. The trained classifier was used to label testing data into one of the two classes. Each classifier predicted the class of the test data, and the overall classification was determined using a majority vote (Polikar 2006), a decision scheme in which the final label is taken to be the one predicted by the majority of classifiers. For example, if the decisions of the three classifiers are respectively, class 1, class 2 and class 1, the majority vote would yield class 1 as the predicted label. The data used for training and testing the classifier is described in Sec. 4.3.6.1. For more information on classification and LDA, please refer to Ref. (Bishop 2006).

4.3.6 Data analyses

4.3.6.1 Determining accuracies

Classification of NIRS signals can either be performed offline, following the completion of data collection, or online, in real-time, as the data are being collected. In general, the aim of offline classification is to provide an estimate of how a classifier, trained on the data collected, would perform on similar future data. Offline classification also provides the ability to make adjustments to the analysis methods, such as extracting and selecting different features. In contrast, online classification involves training a classifier using previously collected data, and then predicting the class of new data in real-time as the task is being performed. Online classification enables real-time control and can provide the user with immediate feedback.

Offline accuracies were calculated retrospectively after all the offline data had been collected. Specifically, all data collected in session 4 and the first block of session 5 were pooled together, and accuracies were determined using thirty iterations of five-fold cross-validation. Cross-validation is a well-established method for statistically estimating classifier performance, namely, how well the classifier will generalize when presented with previously unseen data (Refaeilzadeh, Tang, and Liu 2009). Specifically, cross-validation involved randomly partitioning the data into five equally sized portions (folds). Next, each fold was used as testing data, while the other four folds were used as training data. The training data were used for feature selection and classifier training, and the testing data were used to estimate classification accuracy. This process was repeated until all folds had been used for testing and five classification accuracies had been obtained. Five-fold cross-validation was then repeated twenty-nine more times, with new, randomly partitioned folds. Finally, the 150 accuracies were averaged to provide an overall mean offline accuracy for each participant (30 iterations x 5 folds = 150 accuracies).

Online accuracies were calculated in real-time. The classifier was trained using all the offline data (session 4 and the first block of session 5), and each new task was classified immediately after being performed. Specifically, in the final 2 blocks of Session 5, the data were classified using an online classifier trained on the offline data (Sarah D. Power, Kushki, and Chau 2012).

4.3.6.2 Comparison of ease-of-use

The ease-of-use ratings for each task was summed across all 18 instances where the task was performed. These sums were then used to rank the tasks based on ease-of-use for that participant, where a rank of 1 represented the hardest task to perform and a rank of 11 represented the easiest task to perform. Finally, the ordinal ease-of-use rankings of the two groups were compared using a two-tailed Mann-Whitney *U*-test. For all statistical tests, normality of the data was confirmed using the Shapiro-Wilk Normality test.

4.3.6.3 Comparison of accuracies

The offline and online accuracies achieved over sessions 4 and 5 by the personalized mental task group were compared to the prescribed mental task group using a two-tailed Student's *t*-test for two independent means ($\alpha = 0.05$). Additionally, the personalized mental task groups offline classification accuracies between the participant's personalized tasks and the state-of-the-art prescribed mental strategies (mental math and rest) at the end of session three were compared using a two-tailed Student's *t*-test for two dependent means ($\alpha = 0.05$). The offline classifications were performed using ten iterations of five-fold cross-validation and using two extracted features from the data collected in sessions one through three, which consisted of eighteen data points per task.

4.3.6.4 Evaluation of WS-Score

In order to verify the suitability of the WS-Score, a Pearson correlation between the WS-Score and the online accuracy achieved by the participants was investigated. The total WS-Score for each participant was determined by the absolute difference between the WS-Scores for the increasing and decreasing tasks after the third session, as shown in Equation 7.

$$Total\_WS_{ID} = \left| WS_I - WS_D \right| \quad (7)$$

where *I* is the increasing task in a given pair and *D* is the decreasing task in a given pair.

4.3.6.5 Evaluation of feedback

At the end of the fifth session, participants were asked to rate on a 7-point Likert-type scale how helpful they found the continuous activation feedback to be, with 1 denoting not helpful and 7 meaning most helpful. The helpfulness of the feedback was compared between the two groups using a two-tailed Student's t-test for two independent means (α = 0.05).

4.3.6.6 Analysis of time windows of selected features

A frequency count of the features selected from the different time windows (0-5s, 0-10s, 0-15s, and 0-17s) was conducted with all the data collected during the offline sessions. This count was completed for the selected eight features, including all nineteen participants, for each of the three classifiers (temporal, spatiotemporal, and temporal combined with spatiotemporal). For feature selection methods, please refer to Sec. 4.3.5.3. In total, 456 features were considered.

4.4 Results

4.4.1 Ease-of-use: personalized vs. prescribed tasks

The perceived ease-of-use of the personalized task group and prescribed task group are shown in Figure 17 (A). A Mann-Whitney U-Test revealed that the overall task ease-of-use of the BCI was significantly higher for the personalized mental task group compared to the prescribed mental task group (z = 2.16, p = 0.0308).

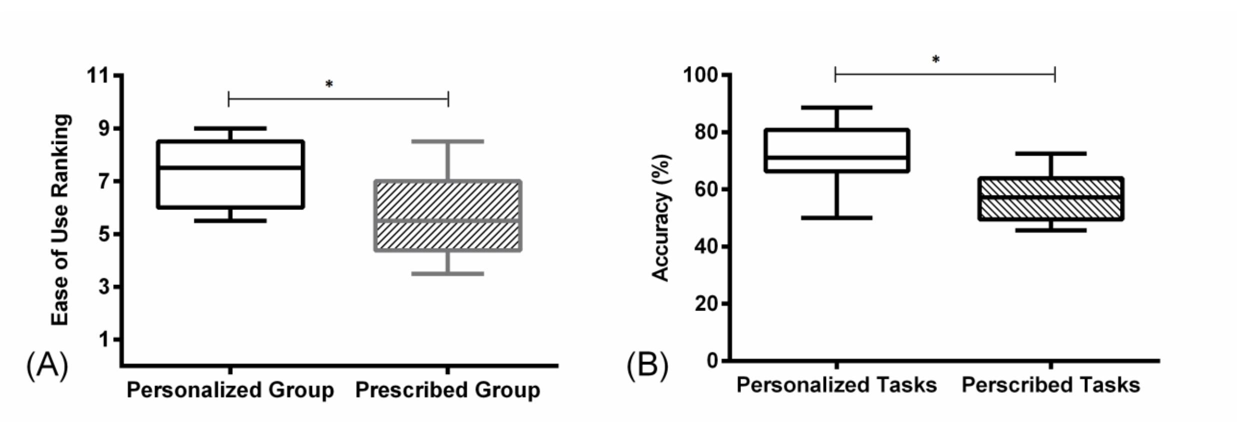

Figure 17. (A) Ease-of-use rankings for personalized and prescribed task groups. (B) Classification accuracy of personalized tasks (chosen after session 3) and prescribed tasks (mental math and rest) for the personalized mental task group in sessions 1-3. Legend: * = p < 0.05.

A high variability was observed in the perceived ease-of-use ratings for tasks that users found to be easiest among both personalized and prescribed task groups. Each task was rated very high (5/5) by some users and very low (1 or 2/5) by other users. The inter-subject variability in the tasks' ease-of-use supports the notion that different individuals find different mental tasks easy to use.

4.4.2 Offline and online classification accuracies

The offline and online classification accuracies achieved in sessions 4 and 5 by the personalized mental task group and the prescribed mental task group are shown in Table 10 and Table 11. Both groups achieved average online classification accuracies greater than 70%, which has been cited as the accuracy required for an effective BCI (Andrea Kübler, Neumann, et al. 2001). However, only the personalized mental task group achieved an average offline accuracy greater than 70%. On average, the personalized mental tasks group achieved an offline accuracy of 75% ± 10.8% and an online accuracy of 77% ± 12.9%, while the prescribed mental task group achieved an offline accuracy of 68% ± 12.9% and an online accuracy of 73% ± 12.9%. Statistically, the classification accuracies achieved by the two groups were not significantly different, as evaluated by a two-tailed *t*-test for two independent means (offline accuracies: $t = 1.29$, $p = 0.213$; online accuracies: $t = 0.554$, $p = 0.587$).

Table 10. Accuracies achieved by prescribed task group.

ID	Offline (%)	Online (%)
101	62.2	52.5
102	76.8	82.5
103	68.2	67.5
104	92.8	95.0
105	57.4	65.0
106	71.6	77.5
107	79.3	82.5
108	47.7	60.0
109	63.2	82.5
110	59.5	67.5
Average	67.9	73.3

Table 11. Accuracies achieved by personalized task group.

ID	Offline (%)	Online (%)
1	81.7	95.0
2	69.2	57.5

3	71.6	77.5
4	76.6	60.0
5	51.4	72.5
7	75.3	72.5
8	90.0	92.5
9	82.5	76.3
10	76.4	85.0
Average	75.0	76.5

For the personalized mental task group, offline analysis was conducted to compare the classification accuracies between the participant's personalized tasks and the state-of-the-art prescribed mental strategies (mental math and rest) at the end of session three. The average offline classification accuracy for the personalized tasks was 71.8 ± 11.5 % versus 57.7 ± 8.8% for prescribed tasks. A Student's *t*-test for two dependent means revealed that the personalized task accuracies were significantly higher than the prescribed task accuracies (t = -2.90, p = 0.0198). These results are shown in Figure 17 (B).

4.4.3 Variability in personalized tasks

The tasks chosen by the participants of the personalized mental task group as their increase and decrease tasks are shown in Figure 18.

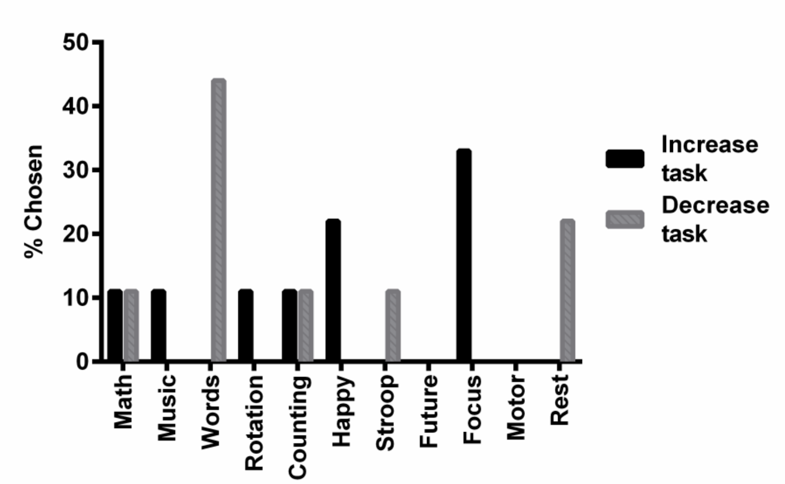

Figure 18. Personalized tasks chosen by users to increase and decrease their haemodynamic activity.

Note that a variety of tasks were chosen for both increase and decrease tasks. Nine of the eleven tasks were chosen at least once, and the only tasks that were not chosen at all were motor imagery and visualizing the future. Of the eleven tasks, relaxing-with-focus was chosen most often as the increasing task, and word generation was most often chosen as the decreasing task. This variability in task choice supports the notion that different individuals prefer different mental tasks. Interestingly, mental math was chosen as an increase task by one participant and a decrease task by another participant. The same phenomenon was observed with the relaxing-with-counting task. This observation suggests a high inter-subject variability in the haemodynamic response produced by each task.

A Hinton plot of the WS-Scores at the end of session three for the personalized mental task group is shown in Figure 19. Inter-subject variability in WS-Scores is evident; each of the eleven tasks resulted in positive WS-Scores in some users and negative scores in others.

As previously mentioned two tasks were not chosen by any of the participants - motor imagery and visualizing the future. Yet, as seen in Figure 19, they were amongst the top three increase and decrease tasks for all participants, except P8. Moreover, motor imagery was the top increase task for P5 and the top decrease task for P10, and visualizing the future was the top increase task for P7. This beckons the question as to why these tasks were never chosen by any of the participants. Upon further analysis of the questionnaires and written comments, many of the participants did not enjoy performing these tasks. They found performing motor imagery to be cumbersome, and visualizing the future to be very abstract and difficult to perform consistently.

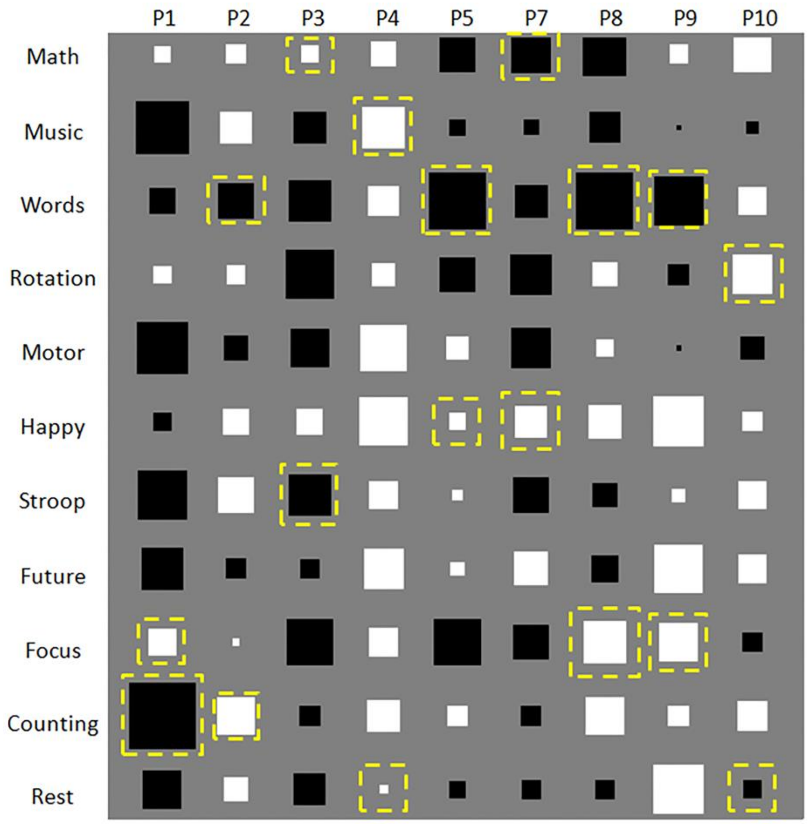

Figure 19. Hinton diagram of WS-Scores at the end of Session 3. Positive and negative values are represented by white and black squares, respectively, and the size of each square is proportional to the magnitude of each WS-Score. Chosen tasks are indicated by a dashed box surrounding the corresponding black or white square. The largest square represents a magnitude of 1.63.

4.4.4 Selection of personalized tasks using the WS-Score

Figure 20 is a scatter plot of each participant's online classification accuracy along with their corresponding total WS-Score.

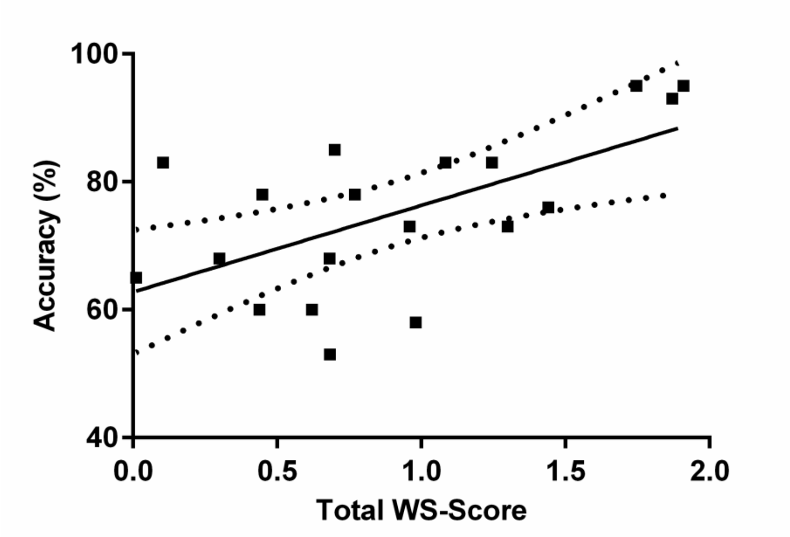

Figure 20. Plot of online accuracy in session 5 versus WS-Score. A regression line was fit to the plot with a slope of 13.54 ± 4.3. The 95% confidence intervals are plotted as dotted lines.

The WS-Scores have a strong positive Pearson's correlation with the online accuracy ($\rho = 0.61$, $p < 0.01$) (Taylor 1990; Hemphill 2003). This suggests that there is potential in using the WS-Score as a measure of task suitability for controlling an NIRS-BCI.

Additionally, we computed the correlation between the offline accuracy achieved in sessions one to three to the total WS-Score at the end of session three over all participants and for all 55 pair wise combinations of tasks. A moderate positive Pearson's correlation was found between the WS-Scores and the offline accuracies ($\rho = 0.4$, $p < 0.001$) (Taylor 1990; Hemphill 2003). This finding reinforces the potential in using the WS-Score as a measure of task suitability for controlling an NIRS-BCI.

4.4.5 Helpfulness of feedback

On average, users found the feedback moderately helpful with a rating of 5.2 ± 1.2 on a 7-point Likert scale. The personalized mental task group had an average helpfulness rating of 5.8 ± 1.0, while the prescribed mental task group had an average helpfulness rating of 4.7 ± 1.1. Based on a Student's *t*-test for two independent means, the personalized mental task group had a significantly higher helpfulness ratings than the prescribed mental task group ($t = 2.27$, $p = 0.036$). Moreover, all participants in the personalized mental task group, other than P7, found the feedback to be helpful (rating > 4). In contrast, only five participants in the prescribed mental task group found the feedback to be helpful (P102, P104, P105, P108, and P109). Overall, this indicates that participants in the personalized mental task group found the feedback significantly more helpful than participants in the prescribed mental task group.

4.4.6 Time window feature selection analysis

All four time windows (0-5s, 0-10s, 0-15s, and 0-17s) were frequently chosen during feature selection for each participant. Overall, the 0-17 second time window was chosen most often, followed by, in descending frequency of selection, the 0-5, 0-15 and 0-10 second time windows, as shown in Figure 21. This result is in line with previous findings by Power et al. (S. Power, Kushki, and Chau 2011). The largest and smallest time windows were chosen most frequently likely because they capture both gradual and early changes in the haemodynamic signal. Additionally, the overall distribution of selected time windows was similar for both the personalized and prescribed mental task groups.

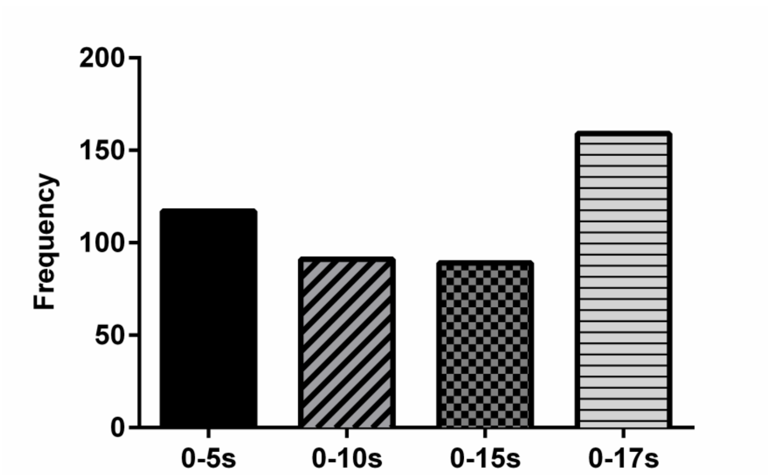

Figure 21. Frequency of occurrence of each time window (0-5s, 0-10s, 0-15s, and 0-17s) among the selected features.

4.5 Discussion

4.5.1 Ease-of-use

Relating to our second research question, the task usability ratings for the personalized task NIRS-BCI were found to be significantly higher than those of the prescribed task NIRS-BCI. This finding is non-trivial because participants selected tasks based on both ease-of-use and WS-scores of each task. Each participant had the opportunity to choose their task by evaluating their own personal ease-of-use/effectiveness tradeoff. Incidentally, our previous offline single-group study also identified a significantly greater perceived ease-of-use for user-selected personalized mental tasks compared to prescribed tasks (Weyand, Takehara-Nishiuchi, and Chau 2015c).

The significance of developing an easier to use BCI has been well established in literature. As described in Sec. 4.2.4, indeed ease-of-use is an identified enabler in the development of BCIs. Ease-of-use is recognized as one of the key attributes to the widespread application of BCI-based communication (J. Wolpaw et al. 2000), one of the most important factor in BCI acceptance (Bos, Poel, and Nijholt 2011; Tan and Nijholt 2010), and one of the most important aspects of the BCI for four severely motor-restricted end-users (Holz et al. 2013). Furthermore, ease-of-use has been linked to satisfaction, which has been shown to positively impact adoption and BCI abandonment (J. Wolpaw et al. 2000; Bos, Poel, and Nijholt 2011; Tan and Nijholt 2010). Overall, an easier to use BCI is vitally important.

4.5.2 Online and offline classification

In support of our first research question, it was determined that individuals can acquire control of an online NIRS-BCI via usability and performance-informed selection of mental tasks while maintaining classification accuracies statistically comparable to those of the prescribed task group. This finding corroborates that of our previous offline single-group study, where no significant difference was observed between the accuracies of user-selected personalized mental tasks and prescribed tasks (Weyand, Takehara-Nishiuchi, and Chau 2015c).

This study adds to the expanding literature of online NIRS-BCI research. Online classification is a critical step towards real-world BCI applications and presents various challenges not applicable to offline classification, including hardware and software adaptations to allow for immediate classification, and to address classifier generalization issues (L. C. Schudlo and Chau 2014). The online accuracies achieved in this study are on par with those reached by Schudlo et al. (L. C. Schudlo and Chau 2014), and Coyle et al. (S. M. Coyle, Ward, and Markham 2007), and exceed the accuracies of other online NIRS-BCI studies, such as those by Chan et al. (Chan, Power, and Chau 2012), and Stangl et al. (Stangl et al. 2013). Our training paradigm was similar to that of previous online NIRS-BCIs (*i.e.* used in Ref. (L. C. Schudlo and Chau 2014)), but with fewer samples for classifier training and a shorter task performance interval of 17 seconds compared to 20 seconds used by Schudlo et al. (L. C. Schudlo and Chau 2014). This shorter response interval can improve the communication rate and decrease mental demand placed on BCI users.

It should be noted, that it is possible that a small non-significant increase in accuracy of personalized mental tasks was actually also present. The power of the online test was calculated to be only 9.8%, and the associated Cohen's *d* effect size was only 0.3 (a small effect). A sample size of 166 participants per group would be necessary to increase the power of this analysis to 80%. Since the effect size appears to be small and a very large number of participants would be required to detect a significant difference, the authors conclude that conducting further analysis using this design is not justified.

However, it is possible that future studies using other tasks or a different length of testing may result in significant differences. For example, with fewer tasks there may be a larger effect size or smaller standard deviation of the personalized mental task group. It is also possible that if longitudinal data were taken for both groups, a greater difference in accuracy may emerge. The ease-of-use of selected tasks may be amplified during extended use, and this could have an effect on the BCI accuracy over time.

Additionally, it should also be noted that users chose their personalized tasks based on both subjective evaluation of performance and usability of the task. Had task choice been exclusively based on performance, a change in accuracy may have been more apparent (Weyand, Takehara-Nishiuchi, and Chau 2015c). However, our findings collectively suggest that perceived ease-of-use may trump accuracy for some users, and may facilitate BCI control. For example, the benefits of personalization in initial acquisition and learning have been demonstrated in other areas of research. In education, personalization has increased learning, motivation, and depth of engagement (Cordova and Lepper 1996). In an air traffic control training study, researchers found that personalized adaptive task selection based on both efficiency and preference led to more efficient training than non-personalized task selection (Salden, Paas, and van Merriënboer 2006).

In line with previous literature, users in the present study achieved significantly higher offline accuracies in some tasks than other tasks (Sorger et al. 2009; Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012). To the best of our knowledge, no other BCI research study has compared online or offline classification accuracies between a personalized and prescribed mental task group. However, one study by Dobrea et al. conducted a within group comparison of personalized mental tasks and prescribed mental tasks. In this EEG-BCI study, Dobrea et al. explicitly compared the offline accuracy of the chosen personalized tasks to a set of prescribed state-of-the-art tasks. Dobrea et al. found that the best combination of four tasks from a choice of twelve tasks (chosen based on classification accuracies), achieved a greater accuracy than the state-of-the-art quartet of mental tasks for all four participants (Dobrea and Dobrea 2009). In line with this result, our study also concluded that the personalized task group achieved significantly higher accuracies using their personalized tasks than the state-of-the-art prescribed tasks (mental math and rest), based on the session three, offline, within-subject classification results.

4.5.3 Variability in haemodynamic changes

Overall, all tasks elicited increases in haemodynamic activity in some participants and decreases in others (Figure 19). The anterior PFC is known to be involved in various executive functions, including: working memory, decision making, predicting future events, multi-tasking, maintaining attention, and emotional control (Gao et al. 1990; Kelly Tai and Chau 2009; Koshino et al. 2011). Additionally, the medial anterior PFC is part of the default mode network (DMN), which is associated with deactivations below resting baseline levels during various goal-directed cognitive tasks and is also activated during autobiographical memory and envisioning the future (Koshino et al. 2011; Buckner, Andrews-Hanna, and Schacter 2008).

The task that resulted in the most consistent increase in haemodynamic activity across participants appears to be happy thoughts. Happy thoughts elicited an increase in activity in all but one participant (P1). This could be due to the fact that happy thoughts involves emotional control, which is believed to be a function of the PFC, and it also involves autobiographical memory, which is known to activate the DMN (Koshino et al. 2011; Buckner, Andrews-Hanna, and Schacter 2008).

Interestingly, the task that appears to result in the most consistent decrease in haemodynamic activity across participants was word generation; it was also the most commonly chosen decrease task. Word generation resulted in a strong decrease in haemodynamic activity in all participants other than P4 and P10. Word generation has often been associated with activations in the left prefrontal cortex (Faress and Chau 2013; Herrmann, Ehlis, and Fallgatter 2003); however, other trends have also been observed (Quaresima et al. 2005). The observed decrease in activation with respect to baseline may be a consequence of measuring mainly over the medial PFC, since the main language areas are predominantly situated on the left side of the brain (Schlösser et al. 1998). Furthermore, the decrease in haemodynamic activity may be attributable to a deactivation in the DMN or resource sharing with the adjacent verbal areas (Shapira-Lichter et al. 2013).

The prescribed tasks of mental math and rest were associated with both activations and deactivations amongst participants. Rest was usually accompanied by decreased haemodynamic activity (all participants other than P2, P4, and P10), while math usually resulted in increased haemodynamic activity (all participants other than P5, P7, and P8). Similar trends have been observed in literature (Bauernfeind et al. 2008; Gert Pfurtscheller, Bauernfeind, et al. 2010; Sarah D. Power, Kushki, and Chau 2012). The increase in haemodynamic activity when performing mental math could be attributed to the engagement of working memory (Sarah Dianne Power and Chau 2013), while the decrease in haemodynamic activity associated with the rest task could be related to mental relaxation. On the other hand, the math task-induced decrease and rest task-associated increase may be related to role of the medial PFC in the DMN (Koshino et al. 2011; Buckner, Andrews-Hanna, and Schacter 2008).

Inter-subject differences in cortical haemodynamic responses may in part be related to inter-individual differences in cognitive processing and brain anatomy. Researchers have shown that there is a large inter-subject variation in the size, shape, and position of various regions of the brain (Nie, Guo, and Liu 2009; Xiong et al. 2000). Thus, it may not be surprising that functional activation of the PFC (the region of focus in our study) varied among participants. EEG-BCI researchers have drawn similar conclusions about the diversity of thought patterns between individuals (Dobrea and Dobrea 2009; Palaniappan 2006).

The large inter-subject variability that appears to be present in most tasks confirms the need for personalized mental strategies. Our results corroborate research showing that the most effective task for controlling a BCI will vary among users (Dobrea and Dobrea 2009; Herff, Heger, Putze, et al. 2013; S. Power, Kushki, and Chau 2011; Ogata, Mukai, and Yagi 2007; Sorger et al. 2009; Nai-Jen and Palaniappan 2004).

4.5.4 Suitability of personalized task selection method

Personalized tasks were chosen on the basis of both performance and ease-of-use. Incidentally, research in human-computer interactions has identified these considerations to be the two most important factors for BCI acceptance (Bos, Poel, and Nijholt 2011).

The WS-Score was proposed as a measure to aid users in choosing their own personalized mental tasks. By providing a method to evaluate each task’s effectiveness, irrespective of task pairings, the WS-Score simplified the selection of personalized mental tasks. Moreover, when using the WS-Score to select personalized tasks, the user is only concerned with one value per task (the task's effectiveness); in contrast, when using classification accuracies, the user is overwhelmed with all 55 pair-wise classification accuracies. The positive correlation between the WS-Score and accuracy (Figure 20) supports the use of the WS-Score as a measure of task effectiveness.

4.5.5 Helpfulness of feedback

On average, the personalized and prescribed task groups found the continuous activation feedback somewhat helpful. This is in line with the findings of a previous NIRS-BCI study by Schudlo et al. where a similar form of feedback was deployed and found to be moderately helpful (3.13 ± 1.25 on a 5 point Likert scale) by users (L. C. Schudlo and Chau 2014).

The personalized mental task group found the feedback to be significantly more useful than did the prescribed mental task group. This could be due to the fact that users in the former group chose their tasks based in part on the feedback. Continuous rather than intermittent (*e.g.* score feedback) feedback may better support long term use of the BCI. Specifically, continuous feedback may promote adaptation of mental strategies and could potentially increase the accuracy and usability of the BCI over time (Niels Birbaumer et al. 2009; Elisabeth V. C. Friedrich, Neuper, and Scherer 2013).

4.5.6 Significance

To date, personalized mental tasks have been explored in MRI (Sorger et al. 2009) and EEG (Dobrea and Dobrea 2009; Palaniappan 2006; Chai et al. 2012) BCIs. To the best of our knowledge, these studies have only explored researcher-selected tasks based solely on performance, with the aim of improving BCI accuracy. In contrast, the present study investigated user-selected personalized tasks with the aim of improving ease-of-use. To the best of our knowledge, to date, the exploration of personalized mental tasks in NIRS-BCIs is limited to one offline, single-group study that illustrated the potential of user-selected tasks in increasing ease-of-use (Weyand, Takehara-Nishiuchi, and Chau 2015c). The present study extends that earlier work by evaluating user-selected personalized mental tasks online in a two-group design.

4.5.7 Limitations and future work

The study was conducted exclusively with able-bodied participants. The findings reported herein likely do not reflect the performance of individuals with severe motor impairments. Further, it would be challenging to perform the personalized mental task protocol with individuals who have reached the total locked-in stage (Niels Birbaumer et al. 2009). Nonetheless, with minor adjustments, we anticipate that the proposed protocol could be applied to clients with incomplete locked-in syndrome who retain reliable visual gaze and a yes/no response. For example, the ease-of-use ratings would need to be administered via a binary selection, scanning paradigm. There are several potential reasons why an individual with locked-in syndrome could stand to benefit from a BCI. Firstly, even if eye gaze has been maintained, muscle fatigue could limit effective communication. Secondly, conditions such as amyotrophic lateral sclerosis are progressive; therefore, when clients transition from a locked-in to a total locked-in state, eye gaze may no longer be a viable access pathway. Literature has suggested that gaining control of the BCI prior to reaching total locked-in syndrome may increase the rate of success (Nikolaus Weiskopf et al. 2007; Niels Birbaumer 2006; Nicolas-Alonso and Gomez-Gil 2012; A. Kübler and Birbaumer 2008). Further research and testing on the target population is necessary before conclusions about the effectiveness of personalized mental tasks in a communication BCI can be drawn.

Secondly, this study was conducted under ideal environmental conditions (quiet and dimly-lit room) that may not be indicative of most real-world settings. Further research should be conducted to assess the effect of environmental conditions on the system's performance.

Finally, when using NIRS as an access modality for a BCI, there is the potential for systemic contributions to the signal (Sarah D. Power, Kushki, and Chau 2012; L. C. Schudlo and Chau 2014; Tachtsidis et al. 2008). Since near-infrared light travels through the scalp and skull before reaching the brain, the recorded signal may contain systemic artefacts. Some researchers have proposed using simultaneous shallow measurements to remove the systemic portion of the deep NIRS signal (Funane et al. 2014; Chan, Power, and Chau 2012). However, the effect of such filtering on classification accuracies has yet to be fully quantified (Chan, Power, and Chau 2012). Other studies by Hoshi et al. and Villinger et al. reported minimal task-related changes in the systemic blood flow (A Villringer et al. 1993; Yoko Hoshi et al. 2011). Furthermore, for the purpose of BCI design, it can be argued that as long as the system is able to differentiate between mental states, the exact origin and composition of the signal may be a moot point.

4.6 Conclusion

This study explored the possibility of allowing participants to choose their own personalized mental tasks, based on both performance and usability, to control an online NIRS-BCI. Our findings suggest that individuals can acquire control of an online personalized NIRS-BCI with classification accuracies comparable to those of an NIRS-BCI with prescribed, state-of-the-art tasks. The personalized mental task NIRS-BCI was significantly easier to use than its prescribed mental task counterpart. Users appeared to be able to effectively choose personalized mental tasks using the WS-Score as the measure of performance, and post-task ease-of-use ratings as the measure of usability. Overall, the personalized mental task NIRS-BCI provided a more user-centered and easier-to-use online BCI, without compromising accuracy. Personalized mental tasks may support the development of more user-friendly BCIs.

Chapter 5: Self-regulation

The entirety of this chapter is reproduced from the article “Weaning off Mental Tasks to Achieve Voluntary Self-Regulatory Control of a Near-Infrared Spectroscopy Brain-Computer Interface”. This manuscript has been published in IEEE Transactions on Neural Systems and Rehabilitation Engineering journal.

5.1 Abstract

As a non-invasive and safe optical measure of haemodynamic brain activity, near-infrared spectroscopy (NIRS) has emerged as a potential brain-computer interface (BCI) access modality. Currently, to the best of our knowledge, all NIRS-BCIs use mental tasks to elicit changes in regional haemodynamic activity. One of the limitations of using mental tasks is that they can be cognitively demanding, and unintuitive. The goal of this work was to explore the development of a neurofeedback-based NIRS-BCI that weans users off mental tasks, to instead use voluntary self-regulation. Ten able-bodied participants were recruited for this study. After ten sessions of using two personalized mental tasks to increase and decrease the participant’s haemodynamic activity, the users were asked, for the remaining sessions, to stop performing their tasks and instead use only a desire to modulate their haemodynamic activity. By the final online session, participants were able to exclusively use voluntary self-regulation with an average accuracy of 79 ± 13%. Additionally, the majority of participants indicated that BCI control via self-regulation was less taxing and more intuitive than BCI operation using mental tasks.

5.2 Introduction

5.2.1 Brain-computer interfaces

Brain-computer interfaces (BCIs) allow individuals to interact with their environment using only cognitive activities (Elisabeth V C Friedrich, Scherer, and Neuper 2012; Ang, Yu, and Guan 2012; S. M. Coyle, Ward, and Markham 2007). BCIs can serve as a conduit to communication or mobility for individuals with severe motor impairments resulting from conditions such as: amyotrophic lateral sclerosis, spinal cord injuries, brain stem stroke, and muscular dystrophy (Elisabeth V C Friedrich, Scherer, and Neuper 2012; Ayaz et al. 2007; Niels Birbaumer 2006; J. Wolpaw et al. 2000; Sitaram et al. 2007). BCIs can also be used by able-bodied individuals for gaming, entertainment, and to accelerate learning (Elisabeth V C Friedrich, Scherer, and Neuper 2012; J. Wolpaw et al. 2000).

The basic components of a BCI are: the physiological input, the signal processing unit, the classifier, and the output. Most BCI research focuses on improving the signal processing unit or the classification methods, with only a few papers focusing on improving the physiological input. The input to the BCI can be categorized into the access modality, which refers to how the physiological signal is collected, and the access pathway, which refers to how a change in the signal is evoked (Blain, Mihailidis, and Chau 2008; K. Tai, Blain, and Chau 2008; Ayaz et al. 2009). The main focus of this research is on improving the BCI access pathway.

5.2.2 Near-infrared spectroscopy access modality

The access modality used in this research is near-infrared spectroscopy (NIRS). NIRS is a non-invasive and safe optical neural imaging technique that measures haemodynamic brain activity (Ang, Yu, and Guan 2012; S. M. Coyle, Ward, and Markham 2007; Ayaz et al. 2007; S. Coyle et al. 2004). Additionally, compared to electroencephalography (EEG), NIRS does not require electrode gel, and is not affected by electrical noise or blinking of the eyes (S. M. Coyle, Ward, and Markham 2007; Ayaz et al. 2009; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; T. Falk et al. 2011; M Izzetoglu et al. 2005). The major limitation of an NIRS-BCI is, however, the inherent haemodynamic delay. For a more detailed analysis of the advantages and limitations of NIRS, please refer to (S. M. Coyle, Ward, and Markham 2007; Ayaz et al. 2009; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; T. Falk et al. 2011; M Izzetoglu et al. 2005).

To make NIRS measurements, a near-infrared light source is placed on the surface of the skin; the light travels through the bone and the meninges to the cortex and is scattered back through the tissue in a banana shaped path, to a detector (S. M. Coyle, Ward, and Markham 2007). The amount of light that is absorbed varies with the amount of oxygen in the blood. Through a mechanism known as neurovascular coupling, areas of the brain that are active typically exhibit an increase in oxygenated haemoglobin (HbO), an increase in total haemoglobin (tHb), and a decrease in deoxygenated haemoglobin (Hb) (S. M. Coyle, Ward, and Markham 2007; Niels Birbaumer and Cohen 2007; Wolf et al. 2002). However, other coupling trends have also been reported (Bauernfeind et al. 2008; A Villringer et al. 1993; Gert Pfurtscheller, Bauernfeind, et al. 2010; Quaresima et al. 2005; Y Hoshi et al. 1994; Koshino et al. 2011; Buckner, Andrews-

Hanna, and Schacter 2008). By measuring the concentration of HbO and Hb in the brain, NIRS provides a measure of cognitive activity (Ang, Yu, and Guan 2012; S. M. Coyle, Ward, and Markham 2007).

5.2.3 NIRS-BCI access pathway: self-regulation

NIRS is a promising access modality; however, to date, little research has investigated the access pathways accompanying this access modality. Currently, to the best of our knowledge, all active NIRS-BCI studies have used mental activation tasks to control the BCI, where the user performs specific tasks that result in predictable changes in haemodynamic activity (Nicolas-Alonso and Gomez-Gil 2012; Strait and Scheutz 2014; L. Schudlo, Weyand, and Chau 2014; Zephaniah and Kim 2014). The main disadvantages of using mental activation tasks are the additional cognitive workload required in performing a separate task, and the non-intuitive nature of this method of communication (Niels Birbaumer 2006; K. Tai, Blain, and Chau 2008). When using mental activation tasks to control a BCI, the user must first determine their intention, then they must recall and perform the task that elicits the corresponding change in haemodynamic activity.

Some BCI research groups are exploring voluntary self-regulation as an alternative BCI access pathway (Niels Birbaumer 2006; N. Weiskopf et al. 2003; Kotchoubey et al. 1996; Andrea Kübler, Neumann, et al. 2001; J. R. Wolpaw et al. 1997; Daly and Wolpaw 2008; A Kübler et al. 1999). Voluntary self-regulation involves the acquisition of voluntary control over one's physiological signals without the need to perform a mental task. One effective procedure for this type of learning is operant conditioning. Operant conditioning was first explored by Edward Thorndike and B.F. Skinner in the late 1800s and involves shaping of a subject's voluntary behaviour by consequences (Coon and Mitterer 2013; Skinner 1948; Thorndike and Rock 1934). Operant conditioning requires two main elements: 1) a subject's voluntary action, and 2) positive or negative reinforcement (Andrea Kübler, Kotchoubey, et al. 2001; Andrea Kübler, Neumann, et al. 2001). Cognitive research has also shown that learning is most effective when the voluntary action causes an immediate, detectable outcome, and when a subject repeatedly undergoes this action-outcome relationship (Pineda et al. 2003). In the application of operant conditioning to a BCI, the action is a subject's voluntary control over a physiological signal, and the outcome is real-time physiological signal feedback. The outcome, depending on whether or not it achieves the intended goal, is followed by selective reinforcement. Operant conditioning allows a subject to acquire the skill of controlling their physiological signals without performing any additional task, and without being consciously aware of how this control is achieved (Niels Birbaumer 2006; N. Weiskopf et al. 2003; Kotchoubey et al. 1996). Operant conditioning promotes the automatization of thought and behaviour, such as in learning how to ride a bike: once learnt, it no longer requires intense concentration or a conscious effort (E. A. Curran and Stokes 2003; Poldrack et al. 2005). This sharply contrasts with mental tasks, which constantly entail an individual's conscious engagement. Thus, self-regulation learned through operant conditioning may require less mental workload than mental tasks.

The field of voluntary self-regulation in BCIs is still in its infancy; however, to date, several researchers have shown the potential of voluntary self-regulation in EEG-BCIs with users gaining control of the 8-12 Hz mu rhythms (J. R. Wolpaw et al. 1997; J R Wolpaw, McFarland, and Vaughan 2000; Daly and Wolpaw 2008; E. A. Curran and Stokes 2003) and slow cortical potentials (SCPs) (Kotchoubey et al. 1996; Andrea Kübler, Neumann, et al. 2001; Daly and Wolpaw 2008; A Kübler et al. 1999; E. A. Curran and Stokes 2003; Niels Birbaumer 2006; N. Birbaumer et al. 1981). While some researchers use a direct approach to learning self-regulation (Kotchoubey et al. 1996; Andrea Kübler, Neumann, et al. 2001; A Kübler et al. 1999; N. Birbaumer et al. 1981), others start by using mental tasks that are gradually replaced by self-regulation (J. R. Wolpaw et al. 1997; Daly and Wolpaw 2008; J R Wolpaw et al. 1991; N. Birbaumer et al. 1999). For example, Wolpaw et al. conducted a study where mu and beta rhythms were used to control vertical cursor movements. Participants first used motor imagery to control vertical cursor movement, but as training progressed, the imagery task was no longer necessary (J. R. Wolpaw et al. 1997; J R Wolpaw et al. 1991). Likewise, Daly et al. also found that many users indicated that they no longer needed their mental tasks after several sessions of training (Daly and Wolpaw 2008).

To date, NIRS self-regulation is relatively unexplored. Toomim et al. (Toomim et al. 2005) and Carmen (Carmen 2005) have conducted studies that show that users are able to voluntarily increase NIRS blood oxygen content in the prefrontal cortex, using feedback without a mental task, for the purpose of treating attention deficit disorder and migraines. However, to the best of our knowledge, no studies have explored voluntary self-regulation for the purpose of controlling an NIRS-BCI, and no NIRS studies have explored the potential of weaning users from their mental tasks.

5.2.4 Objectives

The overall aim of this study was to determine if NIRS-BCI users could be weaned off mental tasks and use just a voluntary desire to alter their cerebral blood oxygenation. The specific objectives were: (i) to determine if a NIRS-BCI based on self-regulation can achieve accuracies greater than chance, and (ii) to determine if users find an NIRS-BCI controlled using self-regulation to be more intuitive, easier to use, and less mentally demanding than a NIRS-BCI controlled using mental tasks.

It was hypothesized that after using mental tasks to increase and decrease their haemodynamic activities, users can be weaned off mental tasks and still be able to increase and decrease their haemodynamic activity. It was also hypothesized that if users can directly modulate their haemodynamic activity without the need to perform a separate task, the resulting BCI would be more intuitive, easier to use, and require less mental workload.

5.3 Methods

5.3.1 Participants

Ten able-bodied participants (four male) between the ages of 16 and 40 were recruited from the staff and students at Holland Bloorview Kids Rehabilitation Hospital (Toronto, Canada). The study received ethics approval from the research ethics boards of Holland Bloorview Kids Rehabilitation Hospital and the University of Toronto. All participants provided informed written consent.

Participants had normal or corrected-to-normal vision and had no known degenerative disorders, cardiovascular disorders, motor impairments, trauma-induced brain injuries, drug or alcohol-related conditions, psychiatric conditions, respiratory disorders or metabolic disorders. Participants were asked not to smoke or drink alcoholic or caffeinated beverages three hours prior to each data collection session. All of the participants were right-handed according to the Edinburgh handedness test (Oldfield 1971). It should be noted that one participant (P6, male) dropped out of the study since he was not able to follow the session protocol.

5.3.2 Instrumentation

NIRS data were collected from the prefrontal cortex (PFC) using a multi-channel frequency-domain NIRS system with a sampling rate of 31.25 Hz (Imagent Functional Brain Imaging System from ISS Inc., Champaign, IL (ISS Inc. 2012)). The PFC is the most anterior portion of the brain lying just behind the forehead and is involved in higher brain functions, including logical thinking, planning, and emotion (Ogata, Mukai, and Yagi 2007; Gao et al. 1990).

Five laser diodes each emitting 690 nm and 830 nm light and three photomultiplier tube detectors were attached to a headband. The headband was made out of a rubber polymer (3M 9900 series). The rubber was comfortable on the skin and easily conformed to the shape of the subject's head. Black fabric was sewn on the outside of the headband to create tight pockets for the light sources and detectors. These pockets secured the sources and detectors, ensuring close and stable contact with the head. The headband was centered on the participant's forehead with reference to the nose, and was placed above the eyebrows, as shown in Figure 22 (A).

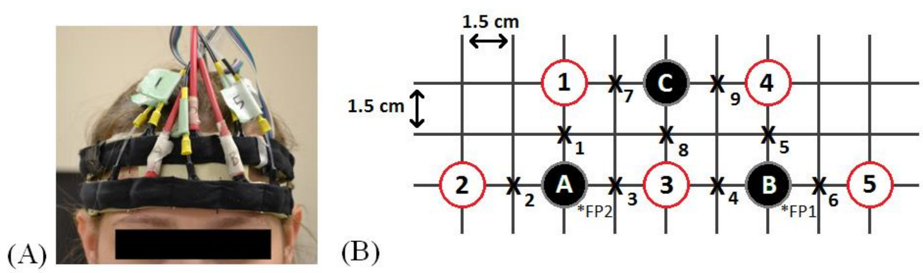

Figure 22. (A) NIRS headband placed over the forehead. (B) Experimental source and detector configuration. Legend: the solid circles represent detectors; the open circles represent light source pairs; the x's represent points of interrogation (channels); and the starred areas represent the approximate FP1 and FP2 positions of the international 10-20 EEG system.

The sources and detectors were arranged in a trapezoidal shape. Each source and adjacent detector was separated by a distance of 3 cm. This distance corresponds to a penetration depth of approximately 2.5 cm, which has been shown to reach the outer layer of the cerebral cortex (Bauernfeind et al. 2008; Haeussinger et al. 2011; E. Okada et al. 1997). Several other NIRS-BCI studies have also used a source-detector separation distance of 3 cm over the PFC (Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; Bauernfeind et al. 2008; Naito et al. 2007; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012). The source-detector configuration allowed for the interrogation of nine discrete locations (channels), as shown in Figure 22 (B).

5.3.3 Experimental protocol

Participants took part in sixteen data collection sessions at a frequency of one session per day. The first fifteen sessions took place at a frequency of five sessions per week for a total of three weeks. The sixteenth session took place ten days after session fifteen. A schematic illustration of the overall study, session, and block structure is shown in Figure 23.

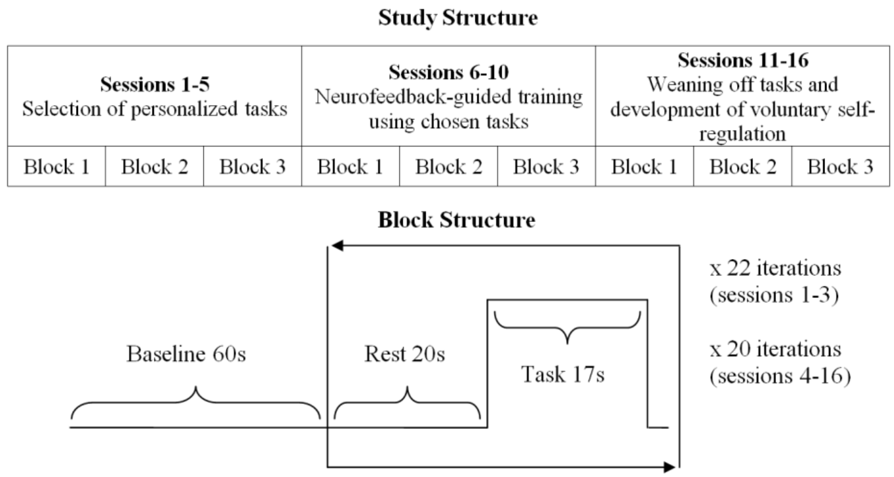

Figure 23. Study, session, and block structure.

All sessions adhered to the same general structure. Each session started with a short warm-up period, which allowed the user to become familiar with the interface. Following the warm up, each participant took part in three data collection blocks. During each data collection block, the participant performed 22 task intervals (sessions 1- 3) or 20 task intervals (sessions 4-16). Each task was performed for 17 seconds, and was followed by a 20 second rest interval. The lengths of the task and rest intervals were chosen on the basis of preliminary data and past work (S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013).

Neurofeedback was provided during all sessions in the form of a graphical user interface consisting of: 1) a trapezoid topographic image showing the real-time changes in blood oxygenation levels over the PFC and 2) a ball that rose and fell with the average change over the entire interrogation area. The feedback was updated every 125 ms, and was calculated using cubic interpolation of the HbO concentrations across the nine points of interrogation (S. M. Coyle, Ward, and Markham 2007; S. Coyle et al. 2004). Participants were informed that the red colour on the feedback display represented an increase in haemodynamic activity, while the blue colour represented a decrease in haemodynamic activity. The activation feedback is shown in Figure 24.

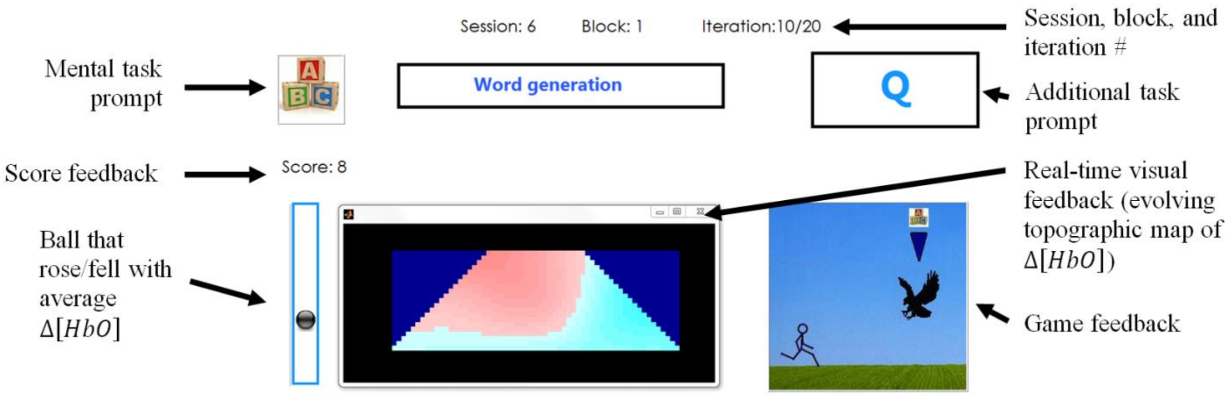

Figure 24. User interface for Sessions 6-10 (online classification and game feedback).

Sessions 1-5: The purpose of the first five sessions was to select personalized mental tasks for each of the participants. In sessions one to three, participants performed eleven mental tasks twice per block in a random order. The eleven mental tasks are described in Table 12.

Table 12. Eleven mental tasks used in sessions 1 to 3.

Task	Description
Mental math (Ang, Yu, and Guan 2012; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Ogata, Mukai, and Yagi 2007; Bauernfeind et al. 2008; Sarah D. Power, Kushki, and Chau 2012; Utsugi et al. 2007; L. C. Schudlo and Chau 2014; Naseer, Hong, and Hong 2014; Khan, Hong, and Hong 2014)	Participants were prompted with a math problem and they were asked to repeatedly subtract a two digit number from a three digit number. For example, given 986-12, the participant would mentally evaluate 986-12 = 974; 974-12 = 962; 962-12 = 950; and so on. Numbers were randomly generated.
Mental singing (Naito et al. 2007; Herff, Heger, Putze, et al. 2013; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011)	Participants were asked to sing a song in their head.
Word generation (Ogata, Mukai, and Yagi 2007; Utsugi et al. 2007; Faress and Chau 2013)	Participants were asked to think of as many words as possible that start with a specific letter. For example, if the letter "D" appeared on the screen, the user may think of the words: dog, door, deli, and so on. Letters (excluding x and z) were randomly generated.
Tangram puzzle (Ayaz et al. 2012; Herff, Heger, Putze, et al. 2013)	Participants were prompted with a tangram puzzle in the top right corner of the screen, and were asked to imagine rotating the pieces to make a final picture.
Counting (Khan, Hong, and Hong 2014; Naseer and Hong	Participants were asked to slowly count in their heads while relaxing.

2013a)	
Happy thoughts (Kelly Tai and Chau 2009)	Participants were asked to think about the details of a past event in their life that made them very happy.
Stroop (Schroeter et al. 2002; Ehlis et al. 2005)	Participants were prompted with a series of colour names. The words were also coloured, but the colour of the words did not always match the written word. For example, the word blue may have been written in red ink. The participants were asked to say the real colour of the word in their head.
Visualizing the future (Buckner, Andrews-Hanna, and Schacter 2008)	Participants were asked to imagine their life in five years, specifically focusing on their future day-to-day activities.
Focus on the feedback (K. Izzetoglu et al. 2011)	Participants were asked to relax and focus on the feedback.
Motor imagery (Kanthack, Bigliassi, and Altimari 2013; Leff et al. 2011; Sitaram et al. 2007; S. M. Coyle, Ward, and Markham 2007; Naseer and Hong 2013b)	Participants were asked to imagine moving their arms or legs.
Rest (Ang, Yu, and Guan 2012; Ayaz et al. 2007; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012; Naseer and Hong 2013a; L. C. Schudlo and Chau 2014; Naseer, Hong, and Hong 2014)	Participants were asked to relax and let their minds rest.

At the beginning of session 4, users were instructed to choose their own pair of personalized tasks based on task performance and their subjective ease-of-use ratings. Task performance was captured by a task-specific weighted slope score that represents the tendency for a task to consistently increase or decrease haemodynamic activity. Specifically, the weighted slope score, WS_i, for the i^{th} task, was defined as:

$$WS_i = \frac{\frac{1}{N}\sum_{j=1}^{N} m_{ij}}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(m_{ij} - \left[\frac{1}{N}\sum_{k=1}^{N} m_{ik}\right]\right)^2}} \quad (8)$$

where m_{ij} and m_{ik} are the slopes of least square line of best-fit to the haemodynamic activity over time for the j^{th} or k^{th} iteration of the i^{th} task, and N is the number of times the task was performed. The ease-of-use ratings were based on post task ease-of-use rankings on a 5-point Likert-type scale, ranging from “very easy” to “very difficult” (Tedesco and Tullis 2006; Sauro and Dumas 2009). A total ease-of-use score for each task was determined as the average ease-of-use rating across all iterations of the task. Each participant chose one task that increased their haemodynamic activity and one task that decreased their haemodynamic activity.

For sessions four and five, participants performed their two chosen mental tasks. It should be noted that once tasks were chosen, participants were not allowed to change their tasks.

Sessions 6-10: For sessions six through ten, participants performed their chosen increase and decrease mental tasks. The goal of these sessions was for the participants to improve at performing their tasks through operant conditioning. All sessions were classified online, in real-time, using a classifier trained on the previously collected personalized mental task sessions (i.e. data from session 4 onwards). In addition to the trapezoid and ball feedback, two additional pieces of feedback were provided in order to motivate participants to perform well. Firstly, the participant was presented with a game involving a character running across a field. The character encountered obstacles that had to be either jumped over or ducked under. If the classification was done correctly, the character would jump or duck as required; however, if the classification was incorrect, the character would run into the obstacle. In addition, a numerical score was displayed to the participant, with a point awarded every time the character successfully avoided an obstacle. A screenshot of the user interface for sessions six through ten is shown in Figure 24.

Sessions 11-15: In sessions 11 to 15, participants were informed that they must stop performing their tasks and instead use voluntary self-regulation. Voluntary self-regulation was described to participants as using a desire to increase or decrease the feedback without performing a separate mental task. They were informed that they should try to decrease and increase the feedback in a manner similar to that achieved with their mental tasks, but without performing the actual mental tasks. Since the transition to self-regulation from mental tasks can be difficult, it was acknowledged that users may want to occasionally revert to mental tasks. Participants were informed that if they felt the need to use some form of their mental task in the first few sessions of self-regulation, this was acceptable. However, participants were reminded that the goal was to gradually minimize the use of tasks with each passing session. Participants were also informed that they would no longer receive prompts for mental tasks and that by the last session, they should not be using their tasks for any iterations.

As participants transitioned from mental tasks to self-regulation, it was hypothesized that the patterns of haemodynamic activity would remain similar, but not necessarily identical. As a result, including data from mental task performance might negatively affect the classification of haemodynamic activity via self-regulation. For this reason, offline self-regulation sessions were required before online self-regulation classification could commence.

To facilitate the training of a new classifier using only self-regulation data, session 11 and the first block of session 12 were analyzed offline. Online classification resumed from session 12, block 2 through session 15, block 3. The online self-regulation classifier was trained using all offline self-regulation sessions (i.e. data from session 11 onwards). The online sessions also included the score and game feedback. The user interface used for session 12, block 2 through session 15, block 3 is shown in Figure 25.

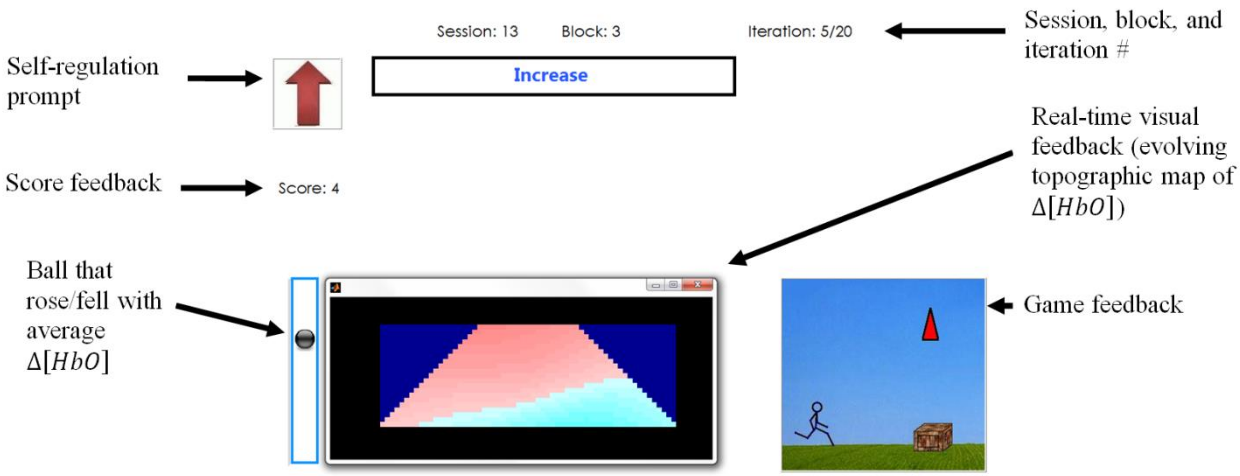

Figure 25. User interface for session 12, block 2 to session 15, block 3 (online classification and game feedback).

Session 16: Session 16 served as a follow-up session and was scheduled approximately ten days after session 15. The goal of session 16 was to determine if participants could still perform self-regulation after a ten day break. The classifier for session 16 was trained on all previously collected self-regulation data (i.e. data from session 11-15), and was only scheduled for participants who achieved an average self-regulation classification accuracy greater than 70%, which has been cited as the accuracy required for an effective BCI (Andrea Kübler, Neumann, et al. 2001).

5.3.4 Data analysis

5.3.4.1 Filtering

The NIRS signal is affected by various sources of physiological noise; including, the Mayer wave oscillation at a frequency of 0.1 Hz, the respiration cycle between 0.2 Hz to 0.4 Hz, and the cardiac cycle between 0.5 Hz to 2 Hz (Ayaz et al. 2009; Sarah D. Power, Falk, and Chau 2010). To remove noise from the signal, the NIRS data were digitally low-pass filtered in real-time using a third-order Chebyshev infinite impulse response (IIR) cascade filter with a pass-band edge frequency of 0.1 Hz, a stop-band edge frequency of 0.5 Hz, and a pass band ripple of 0.1 dB.

5.3.4.2 Calculating haemoglobin concentrations

After filtering the data, the relative change in concentration of Hb, HbO, and the tHb were calculated using the modified Beer-Lambert's Law (S. M. Coyle, Ward, and Markham 2007; J. Wolpaw et al. 2000; Niels Birbaumer and Cohen 2007).

$$\Delta[HbO] = \frac{\varepsilon_{Hb}^{\lambda_2}\left(\frac{\log\left(\frac{I_B^{\lambda_1}}{I_A^{\lambda_1}}\right)}{DPF^{\lambda_1}}\right) - \varepsilon_{Hb}^{\lambda_1}\left(\frac{\log\left(\frac{I_B^{\lambda_2}}{I_A^{\lambda_2}}\right)}{DPF^{\lambda_2}}\right)}{r\left(\varepsilon_{Hb}^{\lambda_2}\varepsilon_{HbO}^{\lambda_1} - \varepsilon_{Hb}^{\lambda_1}\varepsilon_{HbO}^{\lambda_2}\right)} \quad (9)$$

$$\Delta[Hb] = \frac{\varepsilon_{HbO}^{\lambda_2}\left(\frac{\log\left(\frac{I_B^{\lambda_1}}{I_A^{\lambda_1}}\right)}{DPF^{\lambda_1}}\right) - \varepsilon_{HbO}^{\lambda_1}\left(\frac{\log\left(\frac{I_B^{\lambda_2}}{I_A^{\lambda_2}}\right)}{DPF^{\lambda_2}}\right)}{r\left(\varepsilon_{Hb}^{\lambda_1}\varepsilon_{HbO}^{\lambda_2} - \varepsilon_{Hb}^{\lambda_2}\varepsilon_{HbO}^{\lambda_1}\right)} \quad (10)$$

$$\Delta[tHb] = \Delta[Hb] + \Delta[HbO] \quad (11)$$

where I_B^{λ} is the mean light intensity measured at baseline at wavelength λ, I_A^{λ} is the light intensity measured at any given time at wavelength λ, $\varepsilon_{Hb}^{\lambda}$ and $\varepsilon_{HbO}^{\lambda}$ are the specific extinction coefficient of deoxygenated and oxygenated haemoglobin at wavelength λ (for this study: $\varepsilon_{690nm,Hb}$ = 2.1382 $mM^{-1}cm^{-1}$, $\varepsilon_{830nm,Hb}$ = 0.7804 mM^{-1} cm^{-1}, $\varepsilon_{690nm,HbO}$ = 0.3123 mM^{-1} cm^{-1}, and $\varepsilon_{830nm,HbO}$ = 1.0507 mM^{-1} cm^{-1} (Cope 1991)), DPF^{λ} is the differential path factor at wavelength λ (for this study: DPF_{690nm} = 6.51, and DPF_{830nm} = 5.86 (A Duncan et al. 1995)) and r is the geometric distance between the emitter and detector (for this study: r = 3cm).

5.3.4.3 Feature extraction and feature selection

A total of 288 features were extracted from the data. Features were extracted over four time windows (0-5s, 0-10s, 0-15s, and 0-17s). The extracted features included the temporal changes in the three chromophores (Hb, HbO, and tHb) estimated at each of the 9 channels (4 time windows x 3 chromophores x 9 channels = 108 features) and the spatiotemporal features of the zero to fourth order discrete orthogonal Chebyshev image moments (4 time windows x 3 chromophores x 15 image moments = 180 features) (L. C. Schudlo, Power, and Chau 2013). The temporal feature extraction involved normalizing each task interval and then determining the slope over the different time windows using the least-square line of best fit. Before extracting spatial features, topographic images for $\Delta[HbO]$, $\Delta[Hb]$, and $\Delta[tHb]$ were generated by cubic interpolation of the haemoglobin concentration values between locations of empirical integration. After the topographic images were generated, the images were normalized, and the zero to fourth order discrete orthogonal Chebyshev polynomial image moments were extracted from the dynamic topograms to summarize the spatiotemporal features of the data (L. C. Schudlo, Power, and Chau 2013).

After the features were extracted, the sequential forward floating search (SFFS) algorithm was used to select a subset of eight features from the total feature set (L. C. Schudlo, Power, and Chau 2013; Pudil, Novovičová, and Kittler 1994). In this algorithm, the Fisher criterion was used to assess the discriminate capabilities of each feature set, as in (L. C. Schudlo and Chau 2014; L. C. Schudlo, Power, and Chau 2013). A judicious subset of features has been shown to lead to smaller classification errors (Ang, Yu, and Guan 2012). For this study, the target number of features was chosen on the basis of preliminary data and past work (Sarah D Power, Kushki, and Chau 2012; L. C. Schudlo and Chau 2014).

5.3.4.4 Pattern classification

Three linear discriminant analysis classifiers were trained: one using just temporal features, one using just spatiotemporal features, and one using a combination of temporal and spatiotemporal features. All three classifiers predicted the class labels of unseen data, using the majority vote as the predictor (L. C. Schudlo, Power, and Chau 2013). All offline analysis was conducted using 30 iterations of 5-fold cross-validation.

5.3.4.5 Survey

After completing each session, participants filled out a questionnaire on the ease-of-use and intuitiveness of the BCI. The questionnaire consisted of 7-point Likert-type questions that asked the user to evaluate the BCI, as well as compare their current BCI experience with that of previous sessions. Participants also completed the NASA Task Load Index at the end of each session. The NASA Task Load Index is a standardized subjective workload assessment designed for evaluating the mental, physical, and temporal demands of human-machine systems as well as the subject's perceived personal performance, effort and frustration. The score for each factor is weighted by the perceived contribution to total workload and summed to arrive at a final task load index (Hart 2006). The NASA questionnaire has good reliability and validity (Rubio et al. 2004; Hoonakker et al. 2011) and has been used in previous BCI studies (Felton, E.A Williams, J.C Vanderheiden, G.C Radwin 2012; Duvinage et al. 2012).

5.3.4.6 Analysis of selected features

We examined the frequency at which each feature was selected. Specifically, the chosen chromophores (Hb, HbO, and tHb), time windows (0-5s, 0-10s, 0-15s, and 0-17s), and channel locations (Ch1-9) were considered. The chromophore and time window analysis was done on the eight selected features, over all nine participants, for each of the three classifiers (temporal, spatiotemporal, and temporal combined with spatiotemporal), and both classification schemes (mental task and self-regulation). In total, 432 features were analyzed. The channel location analysis was done on the eight selected features, over all nine participants, for the temporal classifier only, and both classification schemes (mental task and self-regulation). In total, 144 features were analyzed.

5.4 Results

5.4.1 Chosen tasks

The personalized tasks chosen by each of the participants during the mental task sessions are shown in Table 13.

Table 13. Increase and decrease task chosen by each participant for the mental tasks sessions. Tasks labelled "(VP)" indicate that this task was associated with a visual prompt.

ID	Increase Task	Decrease Task
1	Relaxing with focus	Relaxing with slow counting
2	Relaxing with slow counting	Word generation (VP)
3	Mental math (VP)	Stroop (VP)
4	Music imagery	Rest
5	Happy thoughts	Word generation (VP)
7	Happy thoughts	Mental math (VP)
8	Relaxing with focus	Word generation (VP)
9	Relaxing with focus	Word generation (VP)
10	Mental rotation (VP)	Rest

5.4.2 Self-regulation accuracies

The average accuracies and 95% confidence intervals achieved during the five sessions of self-regulation (sessions 11-15) over all participants is shown in Figure 26.

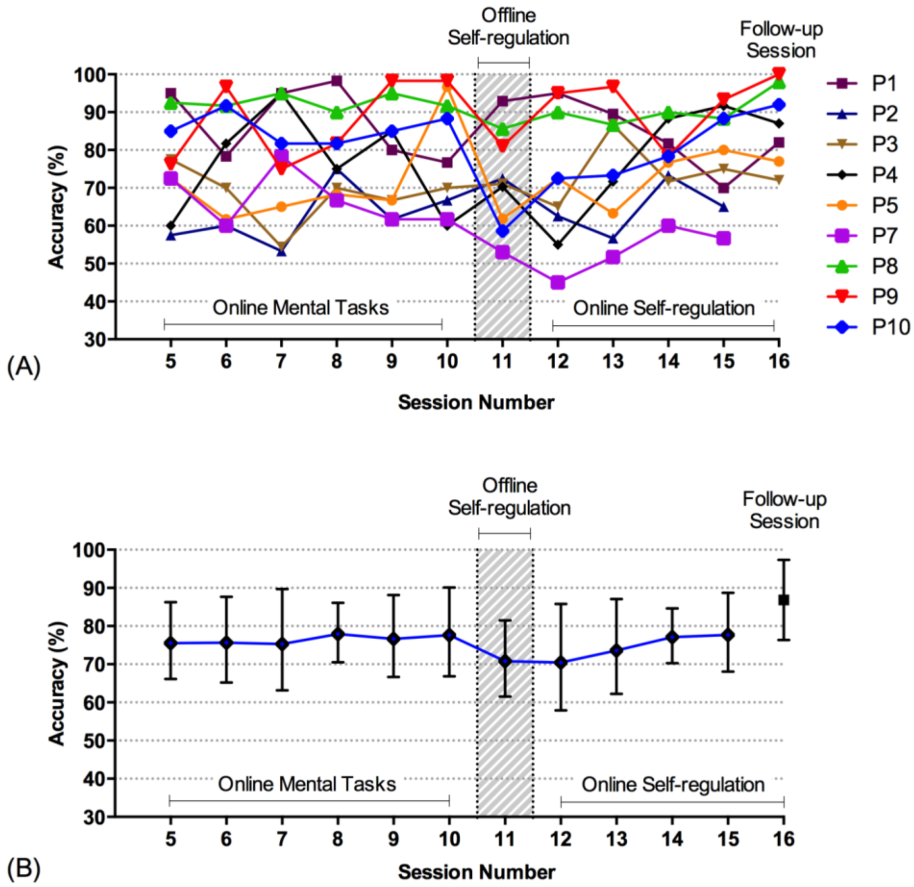

Figure 26. (A) Individual accuracies across all sessions. (B) Average accuracy and 95% CI profile across all participants and all sessions.

The average offline and online accuracies achieved over all participants were 72 ± 13% and 76 ± 14%, respectively, and the average online accuracy achieved in the final online session of self-regulation (session 15) was 79 ± 13%. Other than participant 7, all participants were able to achieve self-regulation accuracies significantly greater than chance, as calculated using the binomial test (two-tailed, $\alpha = 0.05$) (Mueller-Putz et al. 2008).

5.4.3 Follow-up session accuracies

The 7 participants (i.e., all except P2 and P7) who were able to effectively control the BCI using self-regulation in sessions 11-15 (*i.e.* achieved an average accuracy greater than 70% (Andrea Kübler, Neumann, et al. 2001)) took part in an additional session (session 16) approximately 10 days following their last self-regulation session. The summary and spread of the follow-up session accuracy (10 days after session 15) are shown as the final data point in Figure 26.

Normality of the data was confirmed by the Shapiro-Wilk Normality test, and a two-tailed Student's *t*-test (α = 0.05) for two dependent means did not indicate a significant difference between accuracies for sessions 15 and 16 (t = 1.21, p = 0.271). This result suggests that even after a ten day break, participants were still able to control the BCI using voluntary self-regulation without a significant change in accuracy.

5.4.4 Comparison of self-regulation accuracies to mental task accuracies

The accuracies for all six online mental task sessions (sessions 5-10) and four online self-regulation sessions (sessions 12-15) are shown in Figure 26. A linear mixed effect model was run to compare the accuracies obtained using self-regulation to those obtained using mental tasks (West, Welch, and Galecki 2007; Demidenko 2013; Oberg and Mahoney 2007). In the full version of this model, "type of control" (mental tasks or self-regulation) and "session" were modelled as fixed effects, while the intercept and the slope of session over subject were modelled as random effects. The intercept and slope were modelled as random effects to account for inter-subject variations in initial accuracies and changes in accuracies over time due to learning or habituation. The mixed effect model was structured so that "session" was nested in "subject". All model parameters were estimated using maximum likelihood estimation. A *p*-value was obtained by a likelihood ratio test of the full model against the model without the effect in question (method of BCI control). This analysis revealed that the method of BCI control (independent variable) did not have a significant effect on the classification accuracies (χ^2 = 2.31, p = 0.128).

It should be noted that there is an inherent learning bias, which may affect the accuracies achieved. It is possible that if participants had continued using mental tasks instead of transitioning to self-regulation, higher accuracies may have been achieved. However, since self-regulation is meant to follow learning, this bias cannot be easily separated from the analysis. In order to predict whether an increase in accuracy would have likely occurred with continued mental task use, changes in mental task accuracies over the six online mental task sessions were analyzed. A linear mixed effect model was run with "session" as a fixed effect, and the intercept for each subject and the slope of session over subject as random effects. The mixed effect model was structured so that "session" was nested in "subject". All model parameters were estimated using maximum likelihood estimation. The likelihood ratio test revealed no differences in accuracy over time during the mental task sessions (χ^2 = 0.820, p = 0.365). Extrapolating this analysis, it appears that the change in achievable classification accuracies would have been minimal had users continued using mental tasks. Based on this finding, it is likely that the comparison between mental task and self-regulation sessions is minimally affected by the learning bias.

5.4.5 Can mental task data be used to classify self-regulation sessions?

It is possible that the classifier training for self-regulation could be improved or the offline training time reduced by utilizing the training data collected during mental task performance. To determine the suitability of using mental task data for classifying self-regulation sessions, two post-hoc offline data analyses were conducted using a pseudo online approach. First, the classification accuracies achievable using a classifier trained with only mental task data (all data collected during sessions 4-10) and tested on all data collected during each of the participant's self-regulation sessions (sessions 11-15) were evaluated. Second, the classification accuracies achievable when using both mental task and the first block of self-regulation data for classifier training were evaluated. For this latter analysis, the classifier was trained using all mental task data (sessions 4-10) and session 11, block 1 self-regulation data and tested on the remaining self-regulation sessions (session 11, block 2 to session 15). Figure 27 (A) and (B) show the accuracies achieved in the two scenarios mentioned above.

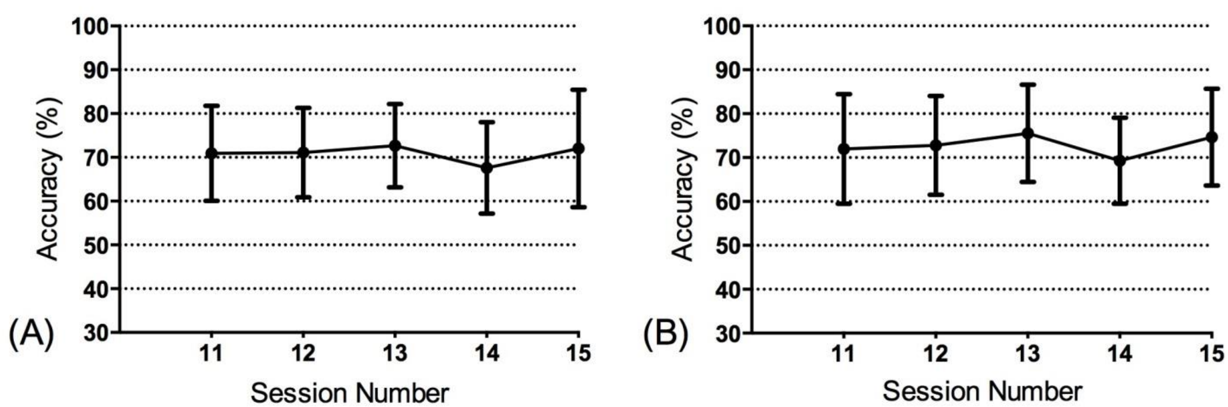

Figure 27. Average accuracy and 95% CI profile for (A) using only mental task data (sessions 4-10) to train the classifier and self-regulation sessions 11 to 15 for testing, and (B) using all mental task data (sessions 4-10) and session 11 block 1 self-regulation data for training the classifier and self-regulation sessions 11 block 2 to 15 for testing.

Under both classification schemes, all participants with the exception of participant 7 achieved an average accuracy greater than chance, as calculated using the binomial test (two-tailed, α = 0.05) (Mueller-Putz et al. 2008). Moreover, the average classification accuracy over all participants for each session, other than session 14, was above 70%. These results suggest that participants were able to elicit very similar changes in their haemodynamic activity during self-regulation as during their mental task sessions.

Two linear mixed effect models were used to compare the accuracies achieved offline (shown in Figure 27 (A) and (B)), to those obtained initially (shown in Figure 26). In the full versions of these models, “training data type” and “session” were modelled as fixed effects, while the intercept and the slope of session over subject were modelled as random effects. In the first model, offline mental task data were compared to the original self-regulation data, while in the second model, offline mental task data and one block of self-regulation data were compared to the original self-regulation data. In both models, the intercept and slope were modelled as random effects to account for inter-subject variations in initial accuracies and changes in accuracies over time due to learning or habituation. The mixed effect models were structured so that “session” was nested in “subject”. All model parameters were estimated using maximum likelihood estimation. A *p*-value was obtained by a likelihood ratio test of each full model against the corresponding model without the effect in question (training data type). The accuracies achieved using only mental task training data were significantly lower than the original accuracies (χ^2 = 5.96, p = 0.0147). However, using both mental task data and one block of self-regulation data for classifier training does not appear to significantly affect the accuracies achieved (χ^2 = 2.25, p = 0.134). For comparison, it’s worth noting that using solely one block of self-regulation data - for example, data from session 11 block 1 - to train a classifier results in chance level accuracies in four of the participants, and an average accuracy of only 64 ± 3%. These results indicate that it may be possible to bridge the change from mental tasks to self-regulation with limited offline self-regulation data collection (i.e. one offline block of self-regulation). However, training with only one block of self-regulation data does not appear to be sufficient.

5.4.6 Are users still performing their tasks?

It appears that the changes in haemodynamic activity elicited during self-regulation sessions are very similar to those elicited during the mental task sessions. Although users were asked to stop performing their mental tasks and instead use voluntary self-regulation, it would be reassuring to confirm that participants did in fact comply. In each post-session questionnaire, participants were asked how frequently they reverted back to performing their mental tasks. The self-reported use of mental tasks is shown in Table 14.

Table 14. Number of times participants reported using mental tasks during self-regulation sessions.

Participant	Session 11	Session 12	Session 13	Session 14	Session 15
P1	0	0	0	0	0
P2	0	2	7	0	0
P3	0	0	0	0	0
P4	0	0	0	0	0
P5	0	0	0	0	0
P7	0	0	0	4	6
P8	0	0	0	0	0
P9	0	0	2	0	0
P10	3	1	0	0	0

Several of the participants (P1, P3, P4, P5, and P8) indicated that they never used their mental tasks during self-regulation sessions. Although a few of the participants (P2, P7, P9, and P10) did perform their mental tasks during self-regulation sessions, none performed their mental tasks for more than 10 of the 60 iterations per session. Additionally, other than participant seven, none of the participants indicated that they performed their mental tasks in session 14 or session 15. Furthermore, it appears that the number of times that participants reverted back to their tasks was not correlated with the initial tasks chosen (Table 13).

From these self-report values and the classification results achieved, it is clear that most participants were able to effectively transition from mental task performance to complete self-regulation while maintaining effective BCI control.

5.4.7 Usability analysis

The NASA Task Load Index for the final session of mental tasks (session 10) and the final session of self-regulation (session 15) are shown in Table 15.

Table 15. NASA Task Load Index during mental tasks and self-regulation.

NASA Task Load Index		
Participant	**Mental Task**	**Self-Regulation**
P1	49.7	39.5

P2	10.5	13.8
P3	55.7	41.5
P4	48.0	16.8
P5	41.2	25.2
P7	40.3	36.2
P8	3.0	0.0
P9	43.3	42.8
P10	17.0	19.0
Average	34.3 ± 19	26.1 ± 15

Seven out of the nine participants (P1, P3, P4, P5, P7, P8, and P9) reported a lower mental workload (i.e. lower NASA task-load score) when using self-regulation in comparison to using mental tasks in the final session of each control method. A two tailed paired *t*-test ($\alpha = 0.05$) to compare the mental workload for the final session of mental tasks and self-regulation was conducted. The self-regulation paradigm resulted in a marginally-significant lower mental workload in comparison to the mental tasks paradigm ($t = -2.24$, $p = 0.0556$). Normality of the data was confirmed using the Shapiro-Wilk Test.

Two linear mixed effect models were used to evaluate the change in subjectively-evaluated mental workload over the course of mental task and self-regulation BCI use. In the full versions of these models, "session" was set as the fixed effect, while the intercept and the slope of session over subject were modelled as random effects. The intercept and slope were modelled as random effects to account for inter-subject variations in initial subjectively-evaluated mental workload and changes in workload over time due to learning. The mixed effect models were structured so that "session" was nested in "subject". All model parameters were estimated using maximum likelihood estimation. A *p*-value was obtained by a likelihood ratio test of each full model against the corresponding model without the effect in question (session). Although the NASA-TLX scores were not significantly different across the six mental task sessions ($\chi^2 = 1.52$, $p = 0.217$), there was a marginally significant decrease in the workload scores during the five self-regulation sessions ($\chi^2 = 3.79$, $p = 0.0517$), suggesting that control became more automatic with self-regulation, but not with mental task performance.

By the last session of self-regulation, seven out of the nine participants reported on the post-session questionnaire, that self-regulation was easier to perform than mental tasks (P2, P4, P5, P7, P8, P9, and P10), and seven out of the nine participants found self-regulation more intuitive than mental tasks (P1, P2, P4, P5, P7, P8, and P10). Furthermore, six out of the nine participants indicated that they would prefer self-regulation over mental tasks (P2, P4, P5, P7, P8, and P10), while one participant did not have a preference (P1).

To determine if there was a change in perceived intuitiveness and ease-of-use for either self-regulation or mental tasks over time, two mixed effect models were run. The full versions of these models used the "session" as the fixed effect, while the intercept and the slope of session over subject were modelled as random effects. The mixed effect models were structured so that "session" was nested in "subject". All model parameters were estimated using maximum likelihood estimation. A *p*-value was obtained by a likelihood ratio test of each full model against the corresponding model without the effect in question (session). Based on the mixed effect models there appears to be a significant increase in the perceived intuitiveness (χ^2 = 6.79, p = 0.00916) and ease-of-use (χ^2 = 6.38, p = 0.0115) of self-regulation over time, but not for mental tasks.

5.4.8 Feature selection analysis

It was found that all four time windows (0-5s, 0-10s, 0-15s, and 0-17s), all three chromophores (Hb, HbO, tHb), and all nine channels were frequently chosen during features selection for each participant. All three chromophores were chosen a comparable number of times, as shown in Figure 28A.

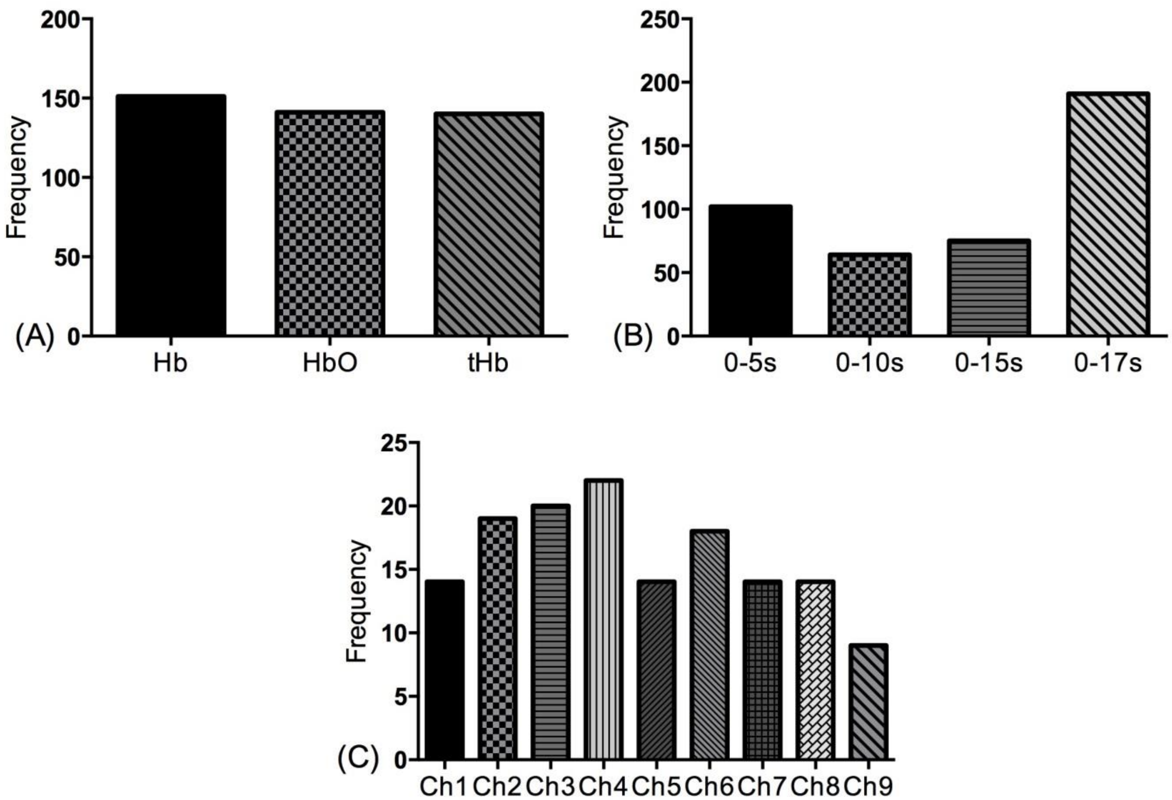

Figure 28. Frequency of occurrence of each (A) chromophore (Hb, HbO, and tHb), (B) time window (0-5s, 0-10s, 0-15s, and 0-17s), and (C) channel, among the selected features. The location of each channel is shown in Figure 22B.

Overall, the 0-17 second time window was chosen most often, followed by, in descending frequency of selection, the 0-5, 0-15 and 0-10 second time windows, as shown in Figure 28B. This finding is aligned with those of Power et al. (S. Power, Kushki, and Chau 2011). The fact that the smallest and largest time windows were chosen most frequently could be attributed to the benefit of capturing both early and gradual changes in the haemodynamic signal.

All channels were chosen frequently, with the channels in the bottom row (Ch2, Ch3, Ch4, and Ch6) being chosen more often than the channels in the middle (Ch1, Ch8, and Ch5) and top (Ch7 and Ch9) rows, as shown in Figure 28C. The location of each channel is shown in the source-detector diagram (Figure 22B). The increased frequency of ventral channel selection could be the result of dorsal-ventral prefrontal cortex variations in organization and activation. Such spatial variations have been observed by several researchers (Wager and Smith 2003; Etkin, Egner, and Kalisch 2011; O’Reilly 2010; Rahm et al. 2013).

Additionally, the trends in feature selection observed above (frequency of chromophores, time windows, and channels chosen) appear to be similar for both mental task and self-regulation sessions.

5.5 Discussion

5.5.1 Accuracy

5.5.1.1 Self-regulation as a means of NIRS-BCI control

With an average online accuracy greater than 70% achieved using self-regulation, the results of this study support that BCI users can indeed be weaned from mental tasks (the typical method of BCI control explored to date) and gain volitional control of their haemodynamic activity. This study adds to the expanding literature of online NIRS-BCI research. The accuracies achieved in this study are in line with those reached by Schudlo et al. (L. C. Schudlo and Chau 2014), and Coyle et al. (S. M. Coyle, Ward, and Markham 2007), and exceed the accuracies of other online NIRS-BCI studies, such as those by Chan et al. (Chan, Power, and Chau 2012), and Stangl et al. (Stangl et al. 2013).

Only one of the 9 participants (P7) was not able to achieve classification accuracies significantly greater than levels of chance using self-regulation. However, given that an estimated 10 to 30% of individuals are BCI-illiterate (Gert Pfurtscheller, Allison, et al. 2010; Ahn et al. 2013), this is perhaps not an alarming finding. Although participant 7 did not achieve an average accuracy greater than chance using self-regulation, he was able to do so using mental tasks. Perhaps this individual needed a different feedback modality (e.g., auditory or tactile), alternative explanations of self-regulation or intermittent rather than continuous feedback. Indeed, contingency, procedural instructions and temporal contiguity have been identified as major factors affecting BCI learning (Sulzer et al. 2013). On the other hand, self-regulation may not be an appropriate method of control for all BCI-users. Methods of systematically selecting the most effective means of control for each NIRS-BCI user should be explored in future work.

5.5.1.2 Self-regulation accuracies after a ten day break

It appears that BCI users with accuracies over 70% are able to take a leave (10 days in this study) from BCI use without a significant decline in accuracy. Since a user may not use the BCI every day, there is value in exploring the stability of accuracy over time. For example, from the perspective of the classification algorithm, there is a possibility that the haemodynamic patterns may change over the ten day break, thereby altering the features and negatively impacting accuracy. Conversely, our results serve as preliminary evidence that haemodynamic activity associated with a particular learned activity does not change greatly over a 10-day period. There is also potential for the user to “unlearn” how to control the BCI, with skill decay after a period of non-use (Arthur et al. 1998). However, our findings suggest that the participants were able to maintain their new skill of self-regulation for at least 10 days. This finding is similar to research showing that learned skills that have reached the stage of automatization are robustly retained and difficult to unlearn (Romano, Howard, and Howard 2010).

5.5.2 Usability

5.5.2.1 Ease-of-use, intuitiveness, and task load

Overall, it appears that a BCI controlled by self-regulation results in a more intuitively-controlled BCI, an easier to use BCI, and a BCI that require less mental workload. All participants, with the exception of P3, found self-regulation superior to mental tasks for at least two of these measures.

The self-regulation BCI appears to be a more flexible BCI, and could lead to greater user satisfaction. Lower effort and mental workload may eventually translate into prolonged, effective and minimally fatiguing BCI use. Various other studies allude to this potential. Sun et al. showed that perceived ease-of-use of an e-learning platform was positively correlated with satisfaction (Sun et al. 2008). Similarly, Roca et al., reported that intention of continuation of learning is positively correlated with satisfaction, which is in turn positively correlated to ease-of-use (Roca, Chiu, and Martínez 2006). Also, self-regulation has been linked to improvements in an individual’s mental health, for example, motor-skill self-regulation was found to be associated with sports confidence (Vealey et al. 1998).

5.5.2.2 User skepticism

Interestingly, most users did not initially find self-regulation more usable than mental tasks. After first being introduced to self-regulation, only one participant (P2) indicated that they would prefer using self-regulation for BCI control. Based on verbal and written comments, several users were initially uncertain that self-regulation could be used effectively. Participant 4 was one of these participants. After achieving an online classification accuracy of 55% in his first online self-regulation session (session 12), this participant indicated that he did not believe self-regulation was possible, and he was unsure about simply expressing a desire to increase or decrease their haemodynamic activity. However, after being encouraged to continue trying, this participant was able to increase his classification accuracies with each session, and changed his attitude towards self-regulation. By the last session of self-regulation (session 15), this participant achieved a 92% classification accuracy and was raving about how much easier and more intuitive self-regulation was in comparison to mental tasks.

Since most of the participants changed their opinion of self-regulation after just a few sessions, the skepticism appears to be short-lived. However, the initial skepticism of using self-regulation for BCI control, which was common among users in this study, could present a barrier for future research exploring self-regulation in BCIs. Additionally, it is possible that user skepticism may also negatively affect self-regulation classification accuracies in the first few sessions. Further research into the effect of user attitude on classification results and how to manage initial skepticism would be beneficial.

5.5.3 Operant conditioning, skill acquisition and learning

Similar to previous EEG-BCI self-regulation results obtained by Wolpaw et al. (J. R. Wolpaw et al. 1997; J R Wolpaw et al. 1991) and Daly et al. (Daly and Wolpaw 2008), this study found that once the skill of controlling the neurofeedback is learned through operant conditioning, the skill becomes automated and mental tasks are no longer needed.

In this study, skill acquisition commenced with resource dependency and ended in automaticity (Langan-fox et al. 2002). Three approximate phases to skill acquisition can be described: the novel phase, where users perform cognitively demanding tasks to achieve changes in the neurofeedback; the weaning phase, where the user reduces his or her reliance on the cognitive processes; and the voluntary control phase, where the user no longer needs to perform their mental tasks. These three phases follow closely the stages of skill acquisition which have been established by numerous researchers in the field of information-processing (Fitts and Posner 1967; Anderson 1982; Shiffrin and Schneider 1977).

While specific brain areas have been implicated in the acquisitoin of skills during certain tasks, such as motor tasks (Poldrack et al. 2005; Gobel, Parrish, and Reber 2011), reading

(Turkeltaub et al. 2003), and attentional expertise (Brefczynski-Lewis et al. 2007), there is a paucity of work on cortical plasticity relating to the acquisition of BCI control.

5.5.4 Significance of study

While self-regulation has been explored in EEG-BCIs (Kotchoubey et al. 1996; Andrea Kübler, Neumann, et al. 2001; J. R. Wolpaw et al. 1997; J R Wolpaw, McFarland, and Vaughan 2000; A Kübler et al. 1999), this is the first study, to the best of our knowledge, to explore the use of self-regulation in an NIRS-BCI. NIRS fundamentally differs from EEG in terms of the measured physiological phenomenon and the method of measurement. Specifically, EEG systems transduce bioelectric manifestations of fast neuronal firing, while NIRS systems optically ascertain haemoglobin chromophore concentration changes related to slow haemodynamic cortical activity. While the bioelectric and haemodynamic phenomena are related by neurovascular coupling, EEG and NIRS methods require different signal processing pathways and generate different neurofeedback. Despite these differences and the haemodynamic delay inherent to NIRS, an average online self-regulation accuracy of 79 ± 13% was achieved by the final session of self-regulation. This accuracy exceeds the often cited threshold of 70% for an effective BCI (Andrea Kübler, Neumann, et al. 2001). Overall, a self-regulation NIRS-BCI appears to offer the same level of accuracy as mental task-driven NIRS-BCIs explored in literature to date, with the added benefits of intuitiveness, ease-of-use, and decreased mental workload.

5.5.5 Study limitations and future directions

This study was conducted over a relatively short period of time, with only five self-regulation sessions. It is possible that individuals gain further proficiency at self-regulation over a longer period of training. Conversely, they might habituate, in which case self-regulation becomes too easy, leading to possible decreases in accuracies. Further research into the long-term use of self-regulation is necessary.

One of the limitations of a self-regulation BCI is that the user must focus some of their attention on the feedback. The attention required could pose difficulties in real-world applications since the user may be focusing on other parts of the interface that they are trying to manipulate. Testing should be conducted to determine how to best design the user interface for a self-regulation BCI. It is important to explore various sizes and on-screen locations of the visual feedback, as well as other types of feedback (*e.g.* audio and tactile). Additionally, the development of a user interface based on a vertically scrolling layout might be well-suited to haemodynamic self-regulation, since moving up and down would be matched well with increasing and decreasing haemodynamic activity.

Finally, this study used a synchronous BCI paradigm that requires the user to always increase or decrease their brain activity. In this sense, a synchronous paradigm can be mentally demanding; future work should explore the possibility of supporting a no-control state to enable a system-paced paradigm (Mason and Birch 2000).

5.6 Conclusions

To the best of our knowledge, this is the first study to explore the use of self-regulation in an NIRS-BCI. Our findings suggest that users can shift from mental task-based to task-free modulation of prefrontal cortical haemodynamics while still maintaining effective BCI control. After 10 sessions of task-based training followed by five sessions of voluntary self-regulation, participants in this study were able to achieve an average online self-regulation classification accuracy of 79 ± 13%. Additionally, participants who achieved an average accuracy over 70% maintained their accuracies after a ten day break from BCI use. Compared to mental tasks, most users found self-regulation to entail a lower mental workload, while being more intuitive and easier to perform. Overall, this research provides an incentive for further exploitation of self-regulation in NIRS-BCI studies.

Chapter 6: Client Study

The entirety of this chapter is reproduced from the article: “Challenges of Implementing a Personalized Mental Task Near-Infrared Spectroscopy Brain-Computer Interface for a Non-Verbal Young Adult with Motor Impairments”. This manuscript has been accepted for publication at Developmental Neurorehabilitation.

6.1 Abstract

Purpose: Near-infrared spectroscopy brain-computer interfaces (NIRS-BCIs) have been proposed as potential motor-free communication pathways. This paper documents the challenges of implementing an NIRS-BCI with a non-verbal, severely and congenitally impaired, but cognitively intact young adult.

Methods: A 5-session personalized mental task NIRS-BCI training paradigm was invoked, whereby participant-specific mental tasks were selected either by the researcher or by the user, on the basis of prior performance or user preference.

Results: Although the personalized mental task selection and training framework had been previously demonstrated with able-bodied participants, the participant was not able to exceed chance level accuracies. Challenges to the acquisition of BCI control may have included disinclination to BCI training, structural or functional brain atypicalities, heightened emotional arousal, and confounding haemodynamic patterns associated with novelty and reward processing.

Conclusions: Overall, we stress the necessity for further clinical NIRS-BCI research involving non-verbal individuals with severe motor impairments.

6.2 Introduction

6.2.1 Near-infrared spectroscopy brain-computer interfaces

Individuals with severe motor impairments, resulting from conditions such as brain-stem injury, spinal cord injury, muscular dystrophy, amyotrophic lateral sclerosis, or other neurological or neuromuscular conditions, often have little or no voluntary muscle control. Despite exhibiting capable cognition, they are unable to express their intentions through conventional means of communication (Blain, Mihailidis, and Chau 2008). The establishment and maintenance of communication has been shown to greatly increase quality of life and autonomy in the face of severe motor impairments (Andrea Kübler, Kotchoubey, et al. 2001; Fenton and Alpert 2008; Bach and McDaniel 1993; Sitaram et al. 2007; Blain, Mihailidis, and Chau 2008; K. Tai, Blain, and Chau 2008).

Brain-computer interfaces (BCIs) enable users to interact with their environment using only cognitive activities (J. Wolpaw et al. 2000; Niels Birbaumer and Cohen 2007). Specifically, near-infrared spectroscopy (NIRS) can be used to measure haemodynamic brain activity and a computer can be trained to discriminate between changes in the hemodynamic response accompanying the performance of different mental tasks. Each task can then be mapped to a command to control a computer. NIRS-BCIs are safe and non-invasive (S. M. Coyle, Ward, and Markham 2007; Ayaz et al. 2007).

6.2.2 NIRS-BCI studies involving individuals with motor impairments

As emphasized by numerous researchers, despite the fact that BCIs have the potential to benefit patients with motor impairments, the vast majority of current research is still being conducted with able-bodied subjects. It is acknowledged that research on able-bodied individuals is important; however, the conclusions may not always transfer to the target population due, for example, to differences in brain structure, cognitive processing pathways, or psycho-behavioural predisposition. Therefore, studies on individuals with motor impairments are critical to further advance clinical BCI’s (Grosse-wentrup and Schölkopf 2013).

Several electroencephalography (EEG) BCI studies have been conducted on individuals with motor impairments with very promising results (N. Birbaumer et al. 1999; A Kübler et al. 1999; G Pfurtscheller et al. 2003; Niels Birbaumer et al. 2000; Jonathan R Wolpaw and McFarland 2004; Andrea Kübler et al. 2009; Andrea Kübler, Neumann, et al. 2001; A Kübler et al. 2005; N.

Birbaumer et al. 1981; Holz et al. 2013). These EEG studies have largely conditioned adult patients to regulate slow cortical potentials or sensory motor rhythms or have exploited event-related potentials. Little research has focused on cognitive task-driven BCIs, an approach that may promote more immediate control, specifically in a pediatric population for whom there is a paucity of EEG BCI research (Moghimi, Kushki, Guerguerian, et al. 2012). Moreover, since NIRS and EEG fundamentally differ in terms of their measured physiological phenomena, the findings from one modality cannot be directly extrapolated to predict the success of the other. To the best of our knowledge, only three NIRS-BCI studies that have been conducted with individuals with motor impairments (Naito et al. 2007; Sarah Dianne Power and Chau 2013; Gallegos-Ayala et al. 2014).

In 2007, Naito et al. conducted a NIRS-BCI study over the prefrontal cortex (PFC) on 40 patients with amyotrophic lateral sclerosis (ALS) between the ages of 22 and 80 years. Seventeen of the subjects had progressed to total locked-in state. Participants performed mental calculation or fast singing as a ‘yes’ response, and number counting, sheep counting, slow singing or landscape imagining as a ‘no’ response. Naito et al. reported an average sensitivity of 75.7% and specificity of 83.5%; however, the study excluded 14 of the 40 participants (35%) because their data were deemed ‘not separable’ based on initial analyses using both the training and testing samples. We note that the numbers of training and testing samples used in this study were very small, with only five data points in each class for testing. Given this small sample size, the upper limit of chance is very high at 76.2% (for a significance threshold of $\alpha = 0.05$ and n = five samples per class) (Mueller-Putz et al. 2008). Based on these calculations, the average sensitivity reported of 75.7% is actually below the upper limit of chance. Other methodological details, such as data splitting into training and testing sets, variations in the number of training samples, and procedure for selecting the increase and decrease tasks for each participant were omitted from the manuscript (Naito et al. 2007).

In 2013, Power et al. conducted an offline NIRS-BCI study over the PFC of a 20 year-old male with Duchenne muscular dystrophy. Using mental arithmetic and a natural baseline state, he achieved an offline classification accuracy of 71.1%, which exceeded chance levels of 63.6%, over two sessions with 24 samples per class in each session. Each session was classified separately and the average of both sessions was presented (Sarah Dianne Power and Chau 2013).

In 2014, Gallegos-Ayala et al. conducted a study on one 67 year-old woman with amyotrophic lateral sclerosis in the completely locked-in stage. Functional activations of the cerebral cortex to auditory processing of correct or incorrect statements were assessed with NIRS. The patient was instructed to think 'yes' or 'no' after each sentence. Response intervals lasted 25 seconds after each question. Online feedback was given as 'your answer was recognized as yes' or 'your answer was recognized as no'. The sensorimotor cortex and temporal areas were interrogated by NIRS. Average prediction accuracies were above chance levels and ranged from 71.67-76.3%. This work serves as preliminary evidence that a locked-in individual can attain control over an NIRS-BCI (Gallegos-Ayala et al. 2014). Despite the promise of these early clinical studies, a broader array of clinical investigations are required to fully ascertain the practical merits and limitations of NIRS-BCIs usage by individuals with various motor impairments.

6.2.3 Objectives

The purpose of this study was to better understand the challenges associated with acquiring control over an NIRS-BCI, i.e. produce machine-discernible haemodynamic changes, in the face of severe congenital motor impairments. The context of our study is a cognitively capable, but severely physically impaired young adult using personalized mental tasks in a five session task selection and training paradigm previously demonstrated in young adults.

6.3 Methods

6.3.1 Participant profile

The participant was a male in his early 20's with an undiagnosed congenital muscular condition that results in severe general hypotonia. The participant has control of eye movements, some facial muscles, as well as very limited leg and upper arm control. All muscle control, including eye movements, deteriorates with fatigue, making muscle-based access difficult. The participant is dependent on a manual wheelchair and is non-verbal, but has reliable 'yes' and 'no' responses using one or two tongue clicks, respectively. The participant also uses a hummer for access to a computer (T. H. Falk et al. 2010). The participant is very bright, and has completed his high-school diploma.

6.3.2 Instrumentation

NIRS data were collected using a multi-channel frequency-domain NIRS system (Imagent Functional Brain Imaging System from ISS Inc., Champaign, IL (ISS Inc. 2012)). The NIRS system was used to measure the blood oxygen content from the PFC (Ogata, Mukai, and Yagi 2007; Gao et al. 1990). A headband with five light sources (emitting 690 nm and 830 nm light) and three photomultiplier tube detectors was centered on the participant's forehead with reference to the nose, and placed directly above the eyebrows, as illustrated in Figure 29. The sources and detectors were separated by a distance of 3 cm and arranged in a trapezoidal shape. With this separation distance, the light has been shown to reach the outer layer of the cerebral cortex (Bauernfeind et al. 2008; Haeussinger et al. 2011; E. Okada et al. 1997). The source-detector configuration allowed for the interrogation of nine discrete locations, between each set of lights sources and detectors (Naito et al. 2007; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Bauernfeind et al. 2008; Sarah D. Power, Kushki, and Chau 2012; Weyand et al. 2015; Weyand, Takehara-Nishiuchi, and Chau 2015a; Weyand, Takehara-Nishiuchi, and Chau 2015b). Each point of interrogation (channel) represents one location where we can estimate the change in the concentration of oxygen in the blood. A schematic diagram of the optode configuration and points of interrogation is shown in Figure 29.

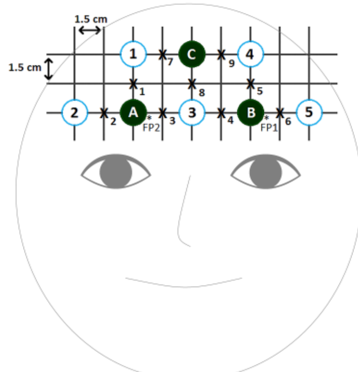

Figure 29. Source-detector placement on forehead. The solid circles represent detectors; the open circles represent light source pairs; the x's represent points of interrogation (channels); and the starred areas represent the approximate FP1 and FP2 positions of the international 10-20 EEG system.

6.3.3 Experimental protocol

The experimental protocol emerged from studies conducted on able-bodied participants. For more information, please refer to (Weyand, Takehara-Nishiuchi, and Chau 2015a; Weyand et al. 2015; Weyand, Takehara-Nishiuchi, and Chau 2015c). Briefly, two personalized task selection approaches were invoked: (1) a user-selected approach that has yielded increased ease-of-use when compared to prescribed tasks (Weyand, Takehara-Nishiuchi, and Chau 2015c; Weyand et al. 2015), and (2) a researcher-selected alternative that has previously led to increased accuracies when compared to prescribed tasks (Weyand, Takehara-Nishiuchi, and Chau 2015c).

The participant took part in five NIRS-BCI sessions, each on a separate day over the span of one and a half months. In each session, the participant was positioned in front of a computer in a dimly lit room. Each session started with a short warm-up period during which the participant familiarized himself with the user interface. Following the warm-up, the participant completed three data collection blocks. During each data collection block, the participant completed 18 (sessions one to three) or 20 (sessions four and five) iterations of mental tasks (described below). Each task was performed for 20 seconds, and was punctuated with a 15-second rest. A schematic of the study, session and block structure is shown in Figure 30.

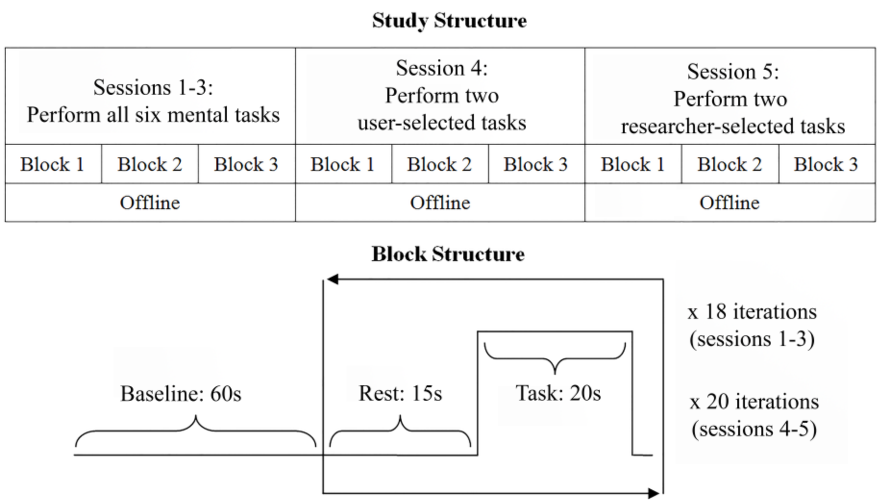

Figure 30. Study, session, and block structure.

6.3.3.1 User interface

Two forms of neurofeedback were provided during each session: 1) a trapezoidal topographic image showing the real-time changes in blood oxygenation levels over the PFC and 2) a ball that rose and fell with the average change over the entire interrogation area. The participant was informed that the red colour on the feedback represented an increase in hemodynamic activity, while the blue colour represented a decrease in hemodynamic activity. For more information on the neurofeedback, please refer to (Weyand et al. 2015; Weyand, Takehara-Nishiuchi, and Chau 2015a; Weyand, Takehara-Nishiuchi, and Chau 2015c). A screenshot of the interface is shown in Figure 31.

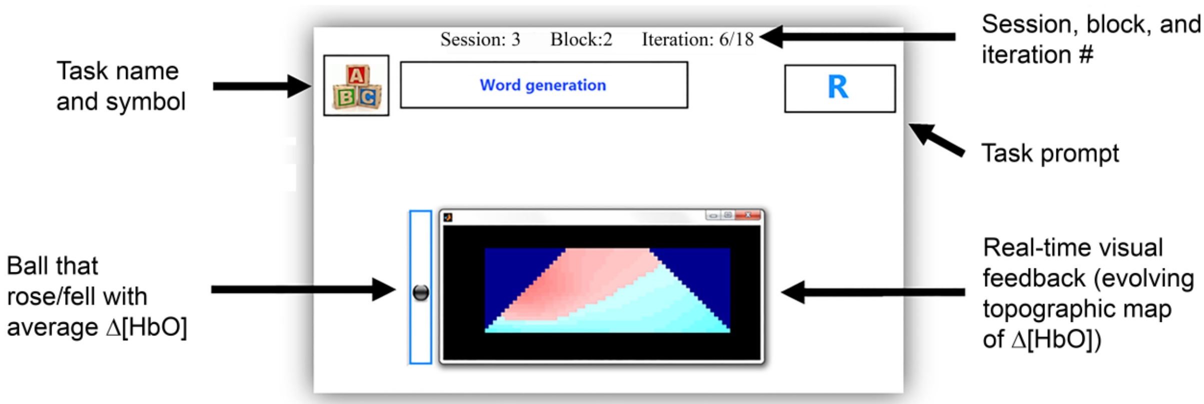

Figure 31. User interface.

6.3.3.2 Session structure

Sessions one to three: The first three sessions were used to collect data on a variety of tasks before personalized tasks were chosen (Weyand, Takehara-Nishiuchi, and Chau 2015a; Weyand, Takehara-Nishiuchi, and Chau 2015c). Specifically, the participant was asked to perform six different mental tasks. The six mental tasks were: mental math, word generation, counting slowly, happy thoughts, focusing on the feedback, and unconstrained rest. Each of the six tasks is briefly described in Table 16; for more details about each task, please refer to (Weyand et al. 2015; Weyand, Takehara-Nishiuchi, and Chau 2015c; Weyand, Takehara-Nishiuchi, and Chau 2015a).

Table 16. Six mental tasks performed by the participant.

Task	Description
Mental math	The participant was asked to repeatedly subtract a one digit number (between two and nine) from a three digit number (Ang, Yu, and Guan 2012; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Ogata, Mukai, and Yagi 2007; Bauernfeind et al. 2008; Sarah D. Power, Kushki, and Chau 2012; Utsugi et al. 2007; L. C. Schudlo and Chau 2014).
Word generation	The participant was asked to think of as many words as possible that start with a prompted letter (Ogata, Mukai, and Yagi 2007; Utsugi et al. 2007; Faress and Chau 2013).
Counting	The participant was asked to slowly count in his head while relaxing (Naseer and Hong 2013a).
Happy thoughts	The participant was asked to think about the details of a past event in his life that made him very happy (Kelly Tai and Chau 2009; Koshino et al. 2011).
Focusing on the feedback	The participant was asked to relax and focus on the feedback (K. Izzetoglu et al. 2011).
Rest	The participant was asked to relax and let his mind wander (Ang, Yu, and Guan 2012; Ayaz et al. 2007; Naito et al. 2007; Herff, Heger, Putze, et al. 2013; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012; Naseer and Hong 2013a; L. C. Schudlo and Chau 2014).

The participant performed each of the six tasks three times per block in a random order. By the end of the third session, the participant had performed each task 27 times (three sessions x three blocks/session x three task iterations/block). At the end of each block, the user was asked to rate each of the six tasks in terms of the ease-of-use and desirability for BCI control on a scale of one to five.

At the end of the third session, the participant chose the mental tasks to be used for controlling the BCI in session four. To inform his choice, the participant was provided with his average ease-of-use rating for each task. In addition, for each task, the participant was presented with two performance metrics: pair-wise accuracy rankings (PWARs) and weighted slope scores (WS-scores). PWARs consisted of a ranked list of task pairs based on their discriminability. The accuracies were calculated after session three using one-iteration of four-fold cross-validation with four selected features for each of the possible pair-wise combinations of the six mental tasks (6 choose 2 = 15). The pairs of tasks were then ranked from highest to lowest. WS-scores measure the tendency of each task to elicit a consistent increase or decrease in hemodynamic activity. WS-scores were calculated by taking the mean of all hemodynamic slopes for a task and dividing by the corresponding standard deviation. The participant was presented with a ranked list of tasks that tended to increase hemodynamic activity and a ranked list of those that tended to decrease hemodynamic activity. For more information on these selection metrics, please refer to (Weyand, Takehara-Nishiuchi, and Chau 2015c).

Session four: In session four, the participant performed the two user-selected tasks chosen at the end of session three, as in (Weyand, Takehara-Nishiuchi, and Chau 2015c; Weyand et al. 2015).

Session five: In session five, the participant performed the two researcher-selected tasks, which were those with the highest PWAR, as in (Weyand, Takehara-Nishiuchi, and Chau 2015c; Weyand, Takehara-Nishiuchi, and Chau 2015a).

6.3.4 Data analysis

6.3.4.1 Accuracies

The NIRS data were filtered using a digital low-pass third-order Chebyshev infinite impulse response (IIR) cascade filter with a pass-band from 0 to 0.1 Hz, a transition band from 0.1 to 0.5 Hz, a stop-band from 0.5 Hz onwards, and a pass band ripple of 0.1. This filter has been used in several previous NIRS BCI studies to mitigate the effect of various sources of physiological noise (Faress and Chau 2013; S. Power, Kushki, and Chau 2011; Sarah Dianne Power and Chau 2013; Weyand, Takehara-Nishiuchi, and Chau 2015c; Weyand et al. 2015; Weyand, Takehara-Nishiuchi, and Chau 2015b; L. C. Schudlo, Power, and Chau 2013; L. C. Schudlo and Chau 2014).

After filtering the data, the changes in concentrations of oxygenated hemoglobin (HbO), deoxygenated hemoglobin (Hb), and total hemoglobin (tHb), were calculated using the modified Beer-Lambert's Law (S. M. Coyle, Ward, and Markham 2007; J. Wolpaw et al. 2000; Niels Birbaumer and Cohen 2007; Kelly Tai and Chau 2009). In this study, the wavelengths, λ, were 690nm and 830nm; the specific extinction coefficients of deoxygenated and oxygenated hemoglobin were $\varepsilon_{690nm,Hb}$ = 2.1382 $mM^{-1}cm^{-1}$ (Cope 1991), $\varepsilon_{830nm,Hb}$ = 0.7804 mM^{-1} cm^{-1} (Cope 1991), $\varepsilon_{690nm,HbO}$ = 0.3123 mM^{-1} cm^{-1} (Cope 1991), $\varepsilon_{830nm,HbO}$ = 1.0507 mM^{-1} cm^{-1} (Cope 1991); the differential path length factors were DPF_{690nm} = 6.51 (A Duncan et al. 1995), and DPF_{830nm} = 5.86 (A Duncan et al. 1995); and the geometric distance between the emitter and detector was r = 3cm.

Next, features were extracted over four time windows (0-5s, 0-10s, 0-15s, and 0-20s). Features included the temporal changes in the three chromophores (Hb, HbO, and tHb) at each of the nine points of interrogation (108 features) and the spatiotemporal features of the zero to fourth order discrete orthogonal Chebyshev image moments (180 features), as in (L. C. Schudlo, Power, and Chau 2013; Weyand, Takehara-Nishiuchi, and Chau 2015a; Weyand, Takehara-Nishiuchi, and Chau 2015b; Weyand et al. 2015; Weyand, Takehara-Nishiuchi, and Chau 2015c). A fast correlation based filter (FCBF) was used to select a subset of eight features from the total feature set for classifier training (Weyand, Takehara-Nishiuchi, and Chau 2015a; Weyand, Takehara-Nishiuchi, and Chau 2015c; Yu and Liu 2003).

Offline accuracies were estimated using ten iterations of ten-fold cross validation (Refaeilzadeh, Tang, and Liu 2009). Three ensemble linear discriminant classifiers were trained: one exclusively with temporal features, a second exclusively with spatiotemporal features, and a third using a combination of temporal and spatiotemporal features. The majority vote of the classifiers was used as the class prediction (Weyand, Takehara-Nishiuchi, and Chau 2015a; Weyand, Takehara-Nishiuchi, and Chau 2015c).

6.3.4.2 Ease-of-use

The nine ease-of-use ratings were averaged for each task over the first three training sessions.

6.3.4.3 Survey

A post-session questionnaire was completed at the end of each session. The participant evaluated the following subjective statements on a seven-point Likert-type scale ranging from 'Strongly Agree' to 'Strongly Disagree': 'I was tired before the session' (after sessions one to five), 'I found it easy to concentrate during the session' (after sessions one to five), 'I liked having the activation feedback' (after sessions one to three), 'It was easy to perform the six tasks' (after sessions one and two), 'It was easy to pick/perform my first task' (after sessions three to five), 'It was easy to pick/perform my second task' (after sessions three to five), 'The feedback was motivating' (after sessions four and five), and 'The feedback was frustrating' (after sessions four and five).

6.4 Results

6.4.1 Tasks-chosen

Based on PWAR, WS-scores, and average ease-of-use ratings, the user-selected tasks were: mental math and relaxing while focusing on the feedback. Interestingly, these are very similar to the tasks most commonly chosen as prescribed tasks by BCI researchers for able-bodied individuals, i.e. mental math and rest (Sarah D. Power, Falk, and Chau 2010; S. Power, Kushki, and Chau 2011; L. C. Schudlo, Power, and Chau 2013; Sarah D. Power, Kushki, and Chau 2012; Ang, Yu, and Guan 2012; L. C. Schudlo and Chau 2014). Moreover, one of the three previous clinical NIRS-BCI studies by Power et al. (Sarah Dianne Power and Chau 2013), used very similar prescribed tasks (mental math and unconstrained rest).

Based on PWAR, the researcher-selected tasks were: word generation and slow counting. Although these tasks did not result in particularly high accuracies after the first three sessions, they were the highest ranked.

6.4.2 Chance level accuracies

Based on a binomial distribution with $\alpha = 0.05$, number of trials = 60, and two classes, the upper confidence limit of chance-level accuracy for sessions four and five was 62.3%. Neither the user-selected tasks (session four; 45.9%) nor the researcher-selected tasks (session five; 54.6%) yielded accuracies significantly greater than chance.

6.4.3 Ease-of-use

The average ease-of-use rating for each task is shown in Table 17. It is evident that the participant enjoyed some tasks more than others. Specifically, the participant favored performing happy thoughts and relaxing while focusing on the feedback. On the other hand, the participant strongly disliked performing slow counting and rest.

Table 17. Average ease-of-use ratings for each of the six tasks.

Tasks	Rating
Counting	1.0
Rest	1.8
Mental math	3.6
Word generation	4.1
Focusing on the feedback	4.9
Happy thoughts	5.0

6.4.4 Survey results

The subjective responses of the participant for each of the survey questions is shown in Table 18. Questions that were not posed in a given session are labelled as not applicable (NA).

Table 18. Post-session responses to survey questions. Legend: 1 = strongly agree, 2 = agree, 3 = somewhat agree, 4 = neutral, 5 = somewhat disagree, 6 = disagree, and 7 = strongly disagree. NA = not applicable.

Statement	Session 1	Session 2	Session 3	Session 4	Session 5
I was tired before the session	2	7	2	7	1
I found it easy to concentrate during the session	1	1	1	1	1
I liked having the activation feedback	1	1	1	NA	NA
It was easy to perform the six tasks	2	1	NA	NA	NA
It was easy to pick/perform my first task	NA	NA	1 (picked math)	1 (math)	1 (words)
It was easy to pick/perform my second task	NA	NA	1 (picked focusing)	1 (focusing)	7 (counting)
The feedback was motivating	NA	NA	NA	1	1
The feedback was frustrating	NA	NA	NA	6	7

6.4.5 Study observations

The participant was very enthusiastic about participating in the research study. During the initial explanation of the study, the participant vocalized to show excitement numerous times and delight was evident in his facial expressions. The participant continued to show enthusiasm and excitement throughout the study. He was eager to start all of the subsequent sessions and rarely wanted to take breaks between blocks. Additionally, when the visual feedback revealed large hemodynamic changes, the participant vocalized to show his excitement and would often look at the researcher and smile. Despite being encouraged to remain calm and focused on the computer, vocalizations and gaze shifting occurred several times in each of the blocks during all five sessions. The participant's parent, who was present at every session, also commented on several occasions about the participant's eagerness to participate in the study.

6.5 Discussion

6.5.1 Effectiveness of five-session personalized mental task selection and training

This research is the first NIRS-BCI clinical case study of personalized mental task selection and training for a user with severe motor impairments. The protocol used in this study has been previously demonstrated with able-bodied individuals (Weyand, Takehara-Nishiuchi, and Chau 2015a; Weyand, Takehara-Nishiuchi, and Chau 2015c; Weyand et al. 2015). All of the classification accuracies for the able-bodied participants in these studies (n = 19) exceeded chance levels.

The participant recruited for this study was not able to gain control over the BCI and achieve accuracies greater than chance levels using either user- or researcher-selected tasks in a five session training paradigm. The participant's inability to control the BCI could be the result of numerous factors; however, we postulate four plausible reasons: 1) disinclination to BCI training (often referred to as the users being 'BCI illiterate'; however, we note that this term is not ICF-compliant); 2) structural or functional brain differences as a result of chronic motor impairments; 3) heightened affective arousal; and 4) contaminant PFC haemodynamic patterns associated with novelty or reward induced activation.

6.5.1.1 Disinclination to BCI training

Some literature has reported that an estimated 10 to 30% of individuals cannot acquire control over a BCI via existing paradigms (Gert Pfurtscheller, Allison, et al. 2010; Ahn et al. 2013). The participant in this study may resemble the fourteen in Naito et al. for whom control was not achieved (Naito et al. 2007). The reasons for a negative predisposition to BCI training are not well understood in the literature. We postulate this disinclination in individuals with severe motor impairments could be attributable to unfamiliarity with mastery-oriented environments and a feeling of helplessness. The participant in this study is likely inexperienced with situations where he is able to effectively control his environment. Although this individual has yes/no responses and hummer access to a computer, his experience in modulating his environment is limited and is likely accompanied by frequent incorrect interpretations of his intentions. As a result, the participant may have, over time, gravitated towards the ‘helplessness’ quadrant of the goal orientation-confidence plane (Leung, Brian, and Chau 2013; Koegel and Mentis 1985). In literature, it is suggested that a feeling of helplessness and low motivation may occur when individuals with disabilities are exposed to frequent failures (Koegel and Mentis 1985). This state may have obfuscated the effort required and what it means to control the BCI, i.e., the individual may not have fully appreciated the need to sustain task performance to demonstrate control.

The methods used in this study (mental tasks) are very similar to those in most NIRS-BCI studies. It is possible that a different training scheme, such as self-regulation (Weyand, Takehara-Nishiuchi, and Chau 2015b), or a different access modality, such as EEG, may better facilitate skill acquisition in this individual. Indeed it has been shown that inability to gain control using one BCI access modality does not preclude successful control of other types of BCIs (Weyand, Takehara-Nishiuchi, and Chau 2015b; Gert Pfurtscheller, Allison, et al. 2010).

6.5.1.2 Participant-specific brain differences

Brain differences secondary to chronic motor impairment may have also contributed to the lack of BCI control. Specifically, there may be atypical cortical structures or haemodynamic activities. Despite the fact that individuals with motor impairments are considered cognitively capable and aware, the function and structure of the brain in these individuals may differ from that of most able-bodied individuals (Neary, Snowden, and Mann 2000; Lillo and Hodges 2010; D’Angelo and Bresolin 2006; Quijano-Roy et al. 2006). Although neuromuscular disorders have traditionally been considered to only affect the motor cortex, recent studies of cognition and behaviour indicate that changes in the prefrontal cortex and other areas of the brain are often present in these individuals (Neary, Snowden, and Mann 2000; Lillo and Hodges 2010; D’Angelo and Bresolin 2006; Quijano-Roy et al. 2006). It has been estimated that up to 50% of patients with motor neuron disease develop some degree of frontal dysfunction relating to attention, working memory, letter fluency, and planning (Lillo and Hodges 2010; Neary, Snowden, and Mann 2000). Moreover, Quijano-Roy et al. showed that four out of six patients with congenital muscular dystrophies and intact cognitive function have cortical and sub-cortical atrophy in the frontal lobe (Quijano-Roy et al. 2006). Additionally, these patients were also found to have enlarged frontal sinuses (Quijano-Roy et al. 2006). This could be a particularly damning problem for NIRS measurements over the PFC as it decreases the volume of gray matter penetrable by light (Haeussinger et al. 2011). Perhaps, functional brain imaging may be required to identify candidate cortical regions and associated personalized tasks for non-invasive BCI control. Conceivably, future BCI protocols should be developed exclusively with the clinical population, rather than with able-bodied individuals.

6.5.1.3 Participant excitement and social validation

It is likely that the participant was hyper-aroused by the feedback to the point that his attention to the task was diminished. From the survey responses and observations made during the study, it was apparent that the participant was satisfied with the sessions and perceived his performance to be excellent. For example, he reported the feedback to be very motivating and not frustrating (Table 18). In fact, the participant appeared to interpret any change in the feedback as positively rewarding, and thus perceived control over the feedback even though the low classification accuracies would suggest random behaviour of the feedback. This erroneous interpretation of the feedback may be attributable in part to the pre-study explanations or the participant’s overwhelming desire to achieve control over the computer. It is possible that for individuals with no physical ability to effect change in their environment, any independent control or perception thereof can be extremely exciting and empowering. Conversely, Leung et al. showed that perception of no control over an access technology can be extremely upsetting and demotivating for a participant (Leung, Brian, and Chau 2013). Perhaps more thorough explanation of the feedback or protocol could help mitigate this problem.

Moreover the participant may have been subconsciously seeking social validation, i.e., the researcher’s ‘approval’ for what he was doing. It is evident from the researcher’s observations that despite being instructed to maintain attention on the computer, the participant’s attention often shifted to the researcher. This shift in attention likely distracted the participant from performing the task at hand. Indeed the need for social validation has been shown in numerous literature (Albrecht, Burleson, and Sarason 1992) and appears to be important for individuals with disabilities (McColl and Skinner 1995; Schulz and Decker 1985).

6.5.1.4 Novelty or reward induced activation

Finally, it is possible that prefrontal activation due to perceived novelty or reward contaminated the task-induced haemodynamic activity. Indeed, the PFC has been shown to be activated as a result of novel stimuli (Daffner et al. 2000; Weierich et al. 2010). Perhaps the novelty of the feedback resulted in activations that were greater than those by the tasks.

The PFC has also been shown to be part of the brain's reward network (Aluja et al. 2015; Rogers et al. 1999; Smith et al. 2011). Literature shows greater startle reflex in the PFC for those who have a higher sensitivity to rewards (Aluja et al. 2015). In a study conducted by Rahimi-Golkhandan et al., it was found that children with developmental coordination disorder have a heightened sensitivity for reward compared to matched controls which is hypothesized to be the result of receiving fewer rewards in real-life settings (Rahimi-Golkhandan et al. 2014). Similarly, it is possible that the participant in this study also has a heightened sensitivity to rewards, and therefore reward-related PFC activations may have masked the underlying task-induced activity.

6.5.2 Limitations and future directions

We urge researchers to consider the pitfalls discussed in this paper when conducting further research on implementing NIRS-BCIs with individuals with severe motor impairments. Firstly, to better gauge the user's instantaneous disposition, we suggest the addition of physiological and observational measures of the participant's excitement and arousal throughout the study. Secondly, where feasible, anatomical brain data ought to be collected or referenced to ascertain the presence of frontal lobe cortical atrophy or enlarged sinuses. Thirdly, we recommend conducting studies with more than five sessions and with varied session frequency, task durations, or session length. It is possible that with more sessions, novelty would subside and the participant would be able to focus more intently on the task at hand. Moreover, changes in the frequency of sessions, task durations, and session length could alter fatigue levels (Mak and Wolpaw 2009; Weyand, Takehara-Nishiuchi, and Chau 2015a), mood (Nijboer et al. 2008; Weyand, Takehara-Nishiuchi, and Chau 2015a), and attention (Hammer et al. 2012; Weyand, Takehara-Nishiuchi, and Chau 2015a), which are known to influence BCI performance. Fourthly, this study only used signals from the PFC. It is possible that the incorporation of other brain regions may unveil more discernible task-related activations. Finally, in this study, a limited task set was developed, drawing largely from literature on able-bodied participants.

Conceivably, unconventional tasks that are more truly personalized to one's disability experience may result in more robust haemodynamic changes. Future studies should seek the input of individuals with motor impairments prior to task development.

6.6 Conclusions

The acquisition of NIRS BCI control can be extremely challenging for a severely impaired individual despite capable cognition. Potential barriers include disinclination to BCI training, lack of familiarity with a mastery-oriented environment, structural or functional brain differences, heightened emotional arousal and contaminant prefrontal haemodynamic patterns associated with novelty or reward processing. In this study, a combination of these factors likely limited the NIRS BCI accuracies of a young adult with severe motor impairments to chance levels, despite invoking previously proven user- and researcher-selected task selection paradigms. We call for future research on NIRS-BCI training paradigms with individuals with disabilities.

Chapter 7: Conclusions

7.1 Summary of Contributions

This thesis makes several original contributions to the field of biomedical engineering. Specifically, in this thesis I have:

1. Conducted the first offline NIRS-BCI study that explored user-selected and researcher-selected personalized mental tasks. **Manuscript published in Journal of Neuroscience Methods** (Weyand, Takehara-Nishiuchi, and Chau 2015c)**.**

    a. Demonstrated the benefit of personalized mental task frameworks. Specifically, it was shown that user-selected personalized tasks resulted in an easier to use BCI compared to prescribed tasks, while researcher-selected personalized tasks resulted in increased accuracy compared to prescribed tasks.

    b. Showed the potential of two user-selected personalized mental task frameworks (WS-scores and PWAR). When comparing the personalized task selection methods, it was concluded that the use of PWARs minimizes the amount of data that needs to be collected, while the use of WS-scores maximizes user satisfaction, and minimizes computational time.

    c. Provided further evidence that there is high inter-subject variability in haemodynamic responses and mental task preferences.

2. Determined the average researcher-selected offline cross-validation accuracies for the 2-, 3-, 4-, and 5-class BCIs to be 78.4 ± 5.7%, 60.5 ± 6.6%, 46.7 ± 5.7%, and 37.2 ± 5.4%, respectively. **Manuscript published in Frontiers in Human Neuroscience** (Weyand, Takehara-Nishiuchi, and Chau 2015a)**.**

    a. Showed that two participants were able to exceed an accuracy of 70% for the 3-class problem.

3. Showed that accuracy was strongly positively correlated (Pearson's) with perceived ease of session (ρ = 0.653), ease of concentration (ρ = 0.634), and enjoyment (ρ = 0.550), but strongly negatively correlated with verbal IQ (ρ = -0.749). **Manuscript**

published in Frontiers in Human Neuroscience (Weyand, Takehara-Nishiuchi, and Chau 2015a)**.**

4. Conducted the first online NIRS-BCI study that explored user-selected personalized mental tasks. **Manuscript published in Neurophotonics** (Weyand et al. 2015)**.**

    a. Verified using an online two group experimental design that a user-selected personalized mental task framework provides heightened usability and a more user-centered design, without sacrificing accuracy.

    b. Provided further evidence that there is high inter-subject variability in haemodynamic responses and mental task preferences.

5. Conducted the first study to explore the use of self-regulation in an NIRS-BCI. **Manuscript published in IEEE Transactions on Neural Systems and Rehabilitation Engineering** (Weyand, Takehara-Nishiuchi, and Chau 2015b)**.**

    a. Demonstrated that users can be weaned off mental tasks to achieve voluntary self-regulation.

    b. Achieved an average online self-regulation classification accuracy of 79 ± 13%.

    c. Showed that participants can maintain their accuracies after a ten day break from BCI use.

    d. Showed that most users found self-regulation to entail a lower mental workload, while being more intuitive and easier to use than mental tasks.

6. Documented the challenges of implementing an NIRS-BCI framework with a severely congenitally impaired, but cognitively intact young adult. **Manuscript accepted at Developmental Neurorehabilitation** (Weyand and Chau 2015)**.**

    a. Found that a client with an undiagnosed motor impairment could not achieve accuracies significantly greater than chance levels.

    b. Postulated that chance-level findings could be due to a combination of factors, including: disinclination to BCI training, lack of familiarity with a mastery-oriented environment, structural or functional brain differences, heightened emotional arousal, or contaminant prefrontal haemodynamic patterns associated with novelty or reward processing.

7.2 Future Work

Several suggestions for future work are described within Chapters 2-6. Overall, we believe the most pressing future research is to continue studies involving clients with severe motor impairments. As this is the target population who currently stand to benefit the most from BCI research, studies investigating the use of various frameworks (specifically, the personalized mental task and self-regulation frameworks described within this thesis) with clients with motor impairments are of utmost importance.

Additionally, we suggest future research explore the use of an initial task set of only four mental tasks. In our personalized mental task studies, we started with either eleven or six mental tasks, and then usually chose the best two tasks for each individual. The benefit of starting with fewer tasks is that each task can be performed more times in each session. Specifically, we suggest starting with mental math, word generation, happy thoughts, and focusing on the feedback. These tasks were chosen often as both user-selected and researcher-selected tasks (Table 2, Table 6, Figure 13, and Figure 18), often resulted in consistent and strong changes in haemodynamic activity (Figure 19), and tended to have high ease-of-use ratings (Figure 7 and Table 17).

References

Advokat, Claire, Nancy Eustis, and John Pickering. 2005. “Relationship Between Diagnosis and Disposition of Patients Admitted to a State Psychiatric Hospital.” *Psychiatric Quarterly* 76 (2): 97–106. doi:10.1007/s11089-005-2333-3.

Ahn, Minkyu, Hohyun Cho, Sangtae Ahn, and Sung Chan Jun. 2013. “High Theta and Low Alpha Powers May Be Indicative of BCI-Illiteracy in Motor Imagery.” *PloS One* 8 (11): e80886. doi:10.1371/journal.pone.0080886.

Albrecht, Terrance, Brant Burleson, and Irwin Sarason. 1992. “Meaning and Method in the Study of Communication and Social Support: An Introduction.” *Communication Research* 19: 149–53.

Allison, Brendan, Thorsten Luth, Diana Valbuena, Amir Teymourian, Ivan Volosyak, and Axel Graser. 2010. “BCI Demographics: How Many (and What Kinds Of) People Can Use an SSVEP BCI?” *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society* 18 (2): 107–16. doi:10.1109/TNSRE.2009.2039495.

Allison, BZ, EW Wolpaw, and JR Wolpaw. 2007. “Brain-Computer Interface Systems: Progress and Prospects.” *Expert Review of Medical Devices* 4 (4): 463–74. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Brain-computer+interface+systems:+progress+and+prospects#0.

Aluja, Anton, Angel Blanch, Eduardo Blanco, and Ferran Balada. 2015. “Affective Modulation of the Startle Reflex and the Reinforcement Sensitivity Theory of Personality: The Role of Sensitivity to Reward.” *Physiology & Behavior* 138. Elsevier Inc.: 332–39. doi:10.1016/j.physbeh.2014.09.009.

An, Jinungn, JH JongHyun Lee, and ChangWook CW Ahn. 2013. “An Efficient GP Approach to Recognizing Cognitive Tasks from fNIRS Neural Signals.” *Science China Information Sciences* 56 (10): 1–7. doi:10.1007/s11432-013-5001-8.

Anderson, JR. 1982. “Acquisition of Cognitive Skill.” *Psychological Review* 89 (4): 369–406. http://psycnet.apa.org/journals/rev/89/4/369/.

Ang, KK, Juanhong Yu, and Cuntai Guan. 2012. “Extracting and Selecting Discriminative Features from High Density NIRS-Based BCI for Numerical Cognition.” In *The 2012 International Joint Conference on Neural Networks (IJCNN)*, 10:1–6. doi:10.1109/IJCNN.2012.6252604.

Arthur, W., W. Bennett, P.L. Stanush, and T.L. McNelly. 1998. “Factors That Influence Skill Decay and Retention: A Quantitative Review and Analysis.” *Human Performance* 11 (1): 57–101. http://www.tandfonline.com/doi/abs/10.1207/s15327043hup1101_3.

Ayaz, Hasan, Meltem Izzetoglu, Scott Bunce, Terry Heiman-Patterson, and Banu Onaral. 2007. “Detecting Cognitive Activity Related Hemodynamic Signal for Brain Computer Interface Using Functional near Infrared Spectroscopy.” In *3rd International IEEE/EMBS Conference on Neural Engineering*, 342–45. Ieee. doi:10.1109/CNE.2007.369680.

Ayaz, Hasan, Patricia a Shewokis, Meltem Izzetoğlu, Murat P Çakır, and Banu Onaral. 2012. “Tangram Solved? Prefrontal Cortex Activation Analysis during Geometric Problem Solving.” In *34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2012:4724–27. doi:10.1109/EMBC.2012.6347022.

Ayaz, Hasan, Patricia Shewokis, Scott Bunce, Maria Schultheis, and Banu Onaral. 2009. “Assessment of Cognitive Neural Correlates for a Functional Near Infrared-Based Brain Computer Interface System.” *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience Lecture Notes in Computer Science*, Lecture Notes in Computer Science, 5638. Springer Berlin / Heidelberg: 699–708. http://dx.doi.org/10.1007/978-3-642-02812-0_79.

Bach, C A, and R W McDaniel. 1993. “Quality of Life in Quadriplegic Adults: A Focus Group Study.” *Rehabilitation Nursing: The Official Journal of the Association of Rehabilitation Nurses* 18 (6): 364–67, 374. http://www.ncbi.nlm.nih.gov/pubmed/7938891.

Banerjee, Amitav, U B Chitnis, S L Jadhav, J S Bhawalkar, and S Chaudhury. 2009. "Hypothesis Testing, Type I and Type II Errors." *Industrial Psychiatry Journal* 18 (2): 127–31. doi:10.4103/0972-6748.62274.

Bauernfeind, Günther, Robert Leeb, Selina Christin Wriessnegger, and Gert Pfurtscheller. 2008. "Development, Set-up and First Results for a One-Channel near-Infrared Spectroscopy System." *Biomed Tech* 53 (1): 36–43. doi:10.1515/BMT.2008.005.

Bauernfeind, Günther, Reinhold Scherer, Gert Pfurtscheller, and Christa Neuper. 2011. "Single-Trial Classification of Antagonistic Oxyhemoglobin Responses during Mental Arithmetic." *Medical & Biological Engineering & Computing* 49 (9): 979–84. doi:10.1007/s11517-011-0792-5.

Bingham, NH, and JM Fry. 2010. "Linear Regression." In *Regression Linear Models in Statistics*, 1–32. Springer.

Birbaumer, N., T. Elber, B. Rockstroh, and W. Lutzenberger. 1981. "Biofeedback of Event-Related Slow Potentials of the Brain." *International Journal of Psychology* 16: 389–415.

Birbaumer, N., N Ghanayim, T Hinterberger, I. Iversen, B. Kotchoubey, A. Kuebler, J Perelmouter, E Taub, and H. Flor. 1999. "A Spelling Device for the Paralysed." *Nature* 398 (6725): 297–98. doi:10.1038/18581.

Birbaumer, Niels. 2006. "Breaking the Silence: Brain–computer Interfaces (BCI) for Communication and Motor Control." *Psychophysiology* 43 (6): 517–32. http://onlinelibrary.wiley.com/doi/10.1111/j.1469-8986.2006.00456.x/full.

Birbaumer, Niels, and Leonardo G Cohen. 2007. "Brain-Computer Interfaces: Communication and Restoration of Movement in Paralysis." *The Journal of Physiology* 579 (3): 621–36. doi:10.1113/jphysiol.2006.125633.

Birbaumer, Niels, Andrea Kübler, Nimr Ghanayim, Thilo Hinterberger, Jouri Perelmouter, Jochen Kaiser, Iver Iversen, Boris Kotchoubey, Nicola Neumann, and Herta Flor. 2000.

“The Thought Translation Device (TTD) for Completely Paralyzed Patients.” *IEEE Transactions on Rehabilitation Engineering* 8 (2): 190–93.

Birbaumer, Niels, Ander Ramos Murguialday, Cornelia Weber, and Pedro Montoya. 2009. “Neurofeedback and Brain-Computer Interface Clinical Applications.” *International Review of Neurobiology* 86: 107–17. doi:10.1016/S0074-7742(09)86008-X.

Bishop, Christopher. 2006. *Pattern Recognition and Machine Learning*. Springer.

Blain, Stefanie, Alex Mihailidis, and Tom Chau. 2008. “Assessing the Potential of Electrodermal Activity as an Alternative Access Pathway.” *Medical Engineering & Physics* 30 (4): 498–505. doi:10.1016/j.medengphy.2007.05.015.

Boas, DA, CE Elwell, Marco Ferrari, and Gentaro Taga. 2014. “Twenty Years of Functional near-Infrared Spectroscopy: Introduction for the Special Issue.” *NeuroImage* 85. Elsevier Inc.: 1–5. doi:10.1016/j.neuroimage.2013.11.033.

Bogacz, Rafal, Eric-Jan Wagenmakers, Birte U Forstmann, and Sander Nieuwenhuis. 2010. “The Neural Basis of the Speed-Accuracy Tradeoff.” *Trends in Neurosciences* 33 (1): 10–16. doi:10.1016/j.tins.2009.09.002.

Bos, DPO, Mannes Poel, and Anton Nijholt. 2011. “A Study in User-Centered Design and Evaluation of Mental Tasks for BCI.” *Advances in Multimedia Modeling Lecture Notes in Computer Science* 6524: 122–34. http://link.springer.com/chapter/10.1007/978-3-642-17829-0_12.

Brefczynski-Lewis, J a, A Lutz, H S Schaefer, D B Levinson, and R J Davidson. 2007. “Neural Correlates of Attentional Expertise in Long-Term Meditation Practitioners.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (27): 11483–88. doi:10.1073/pnas.0606552104.

Buckner, Randy L, Jessica R Andrews-Hanna, and Daniel L Schacter. 2008. “The Brain’s Default Network: Anatomy, Function, and Relevance to Disease.” *Annals of the New York Academy of Sciences* 1124: 1–38. doi:10.1196/annals.1440.011.

Carmen, JA. 2005. “Passive Infrared Hemoencephalography: Four Years and 100 Migraines.” *Journal of Neurotherapy* 8 (3): 23–51. doi:10.1300/J184v08n03_03.

Chai, Rifai, S Ling, G.P. Hunter, and H.T. Nguyen. 2012. “Mental Non-Motor Imagery Tasks Classifications of Brain Computer Interface for Wheelchair Commands Using Genetic Algorithm-Based Neural Network.” In *International Joint Conference on Neural Networks (IJCNN)*, 1–7. doi:10.1109/IJCNN.2012.6252499.

Chan, Justin, Sarah Power, and Tom Chau. 2012. “Investigating the Need for Modelling Temporal Dependencies in a Brain-Computer Interface with Real-Time Feedback Based on near Infrared Spectra.” *Journal of Near Infrared Spectroscopy* 20 (1): 107–16. doi:10.1255/jnirs.971.

Chanel, Guillaume, K Ansari-Asl, and Thierry Pun. 2007. “Valence-Arousal Evaluation Using Physiological Signals in an Emotion Recall Paradigm.” In *IEEE International Conference on Systems, Man and Cybernetics*, 41:2662–67. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4413638.

Coon, Dennis, and J Mitterer. 2013. *Psychology: A Journey*. 5th ed. Wadsworth, Cengage Learning. http://books.google.ca/books?id=OQKS2mA5QgYC.

Cope, Mark. 1991. “The Application of near Infrared Spectroscopy to Non Invasive Monitoring of Cerebral Oxygenation in the Newborn Infant.” University of London. http://discovery.ucl.ac.uk/1317956/.

Cordova, D, and M Lepper. 1996. “Intrinsic Motivation and the Process of Learning: Beneficial Effects of Contextulization, Personalization, and Choice.” *Journal of Educational Pyschology* 88 (4): 715–30.

Coyle, Shirley M, Tomás E Ward, and Charles M Markham. 2007. “Brain-Computer Interface Using a Simplified Functional near-Infrared Spectroscopy System.” *Journal of Neural Engineering* 4 (3): 219–26. doi:10.1088/1741-2560/4/3/007.

Coyle, Shirley, Tomás Ward, Charles Markham, and Gary McDarby. 2004. "On the Suitability of near-Infrared (NIR) Systems for next-Generation Brain–computer Interfaces." *Physiological Measurement* 25 (4): 815–22. doi:10.1088/0967-3334/25/4/003.

Cui, Xu, Signe Bray, DM Bryant, GH Glover, and AL Reiss. 2011. "A Quantitative Comparison of NIRS and fMRI across Multiple Cognitive Tasks." *Neuroimage* 54 (4): 2808–21. doi:10.1016/j.neuroimage.2010.10.069.A.

Curran, Eleanor A., and Maria J. Stokes. 2003. "Learning to Control Brain Activity: A Review of the Production and Control of EEG Components for Driving Brain–computer Interface (BCI) Systems." *Brain and Cognition* 51 (3): 326–36. doi:10.1016/S0278-2626(03)00036-8.

Curran, Eleanor, Peter Sykacek, Maria Stokes, Stephen J. Roberts, Will Penny, Ingrid Johnsrude, and Adrian M. Owen. 2004. "Cognitive Tasks for Driving a Brain-Computer Interfacing System: A Pilot Study." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 12 (1): 48–54. doi:10.1109/TNSRE.2003.821372.

D'Angelo, Maria Grazia, and Nereo Bresolin. 2006. "Cognitive Impairment in Neuromuscular Disorders." *Muscle & Nerve* 34 (1): 16–33. doi:10.1002/mus.20535.

Daffner, K R, M M Mesulam, L F Scinto, D Acar, V Calvo, R Faust, A Chabrerie, B Kennedy, and P Holcomb. 2000. "The Central Role of the Prefrontal Cortex in Directing Attention to Novel Events." *Brain : A Journal of Neurology* 123 (Pt 5: 927–39. doi:10.1093/brain/123.5.927.

Daly, JJ, and JR Wolpaw. 2008. "Brain–computer Interfaces in Neurological Rehabilitation." *The Lancet Neurology* 7 (11): 1032–43. doi:10.1016/S1474-4422(08)70223-0.

Delpy, D T, M Cope, P van der Zee, S Arridge, S Wray, and J Wyatt. 1988. "Estimation of Optical Pathlength through Tissue from Direct Time of Flight Measurement." *Physics in Medicine and Biology* 33 (12): 1433–42. http://iopscience.iop.org/0031-9155/33/12/008.

Demidenko, Eugene. 2013. “Introduction: Why Mixed Models?” In *Mixed Models: Theory and Applications with R*, Second Edi, 1–40. John Wiley & Sons, Inc.

Dietterich, TG, and G Bakiri. 1995. “Solving Multiclass Learning Problems via Error-Correcting Output Codes.” *Journal of Artificial Intelligence Research* 2 (1): 263–86. http://arxiv.org/abs/cs/9501101.

Dobrea, M.-C., and D.M. Dobrea. 2009. “The Selection of Proper Discriminative Cognitive Tasks - A Necessary Prerequisite in High-Quality BCI Applications.” In *2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies*, 1–6. doi:10.1109/ISABEL.2009.5373706.

Duncan, A, J H Meek, M Clemence, C E Elwell, L Tyszczuk, M Cope, and D T Delpy. 1995. “Optical Pathlength Measurements on Adult Head, Calf and Forearm and the Head of the Newborn Infant Using Phase Resolved Optical Spectroscopy.” *Physics in Medicine and Biology* 40 (2): 295–304. http://www.ncbi.nlm.nih.gov/pubmed/7708855.

Duncan, Arlene, Judith H Meek, Matthew Clemence, Clare E Elwell, Penny Fallon, Lidia Tyszczuk, Mark Cope, and David T Delpy. 1996. “Measurement of Cranial Optical Path Length as a Function of Age Using Phase Resolved near Infrared Spectroscopy.” *Pediatric Research* 39 (5): 889–94. doi:10.1203/00006450-199605000-00025.

Dupont, William D. 2009. “Simple Linear Regression.” In *Statistical Modeling for Biomedical Researchers A Simple Introduction to the Analysis of Complex Data*, 45–96. Cambridge University Press.

Duvinage, Matthieu, Thierry Castermans, Mathieu Petieau, Karthik Seetharaman, Thomas Hoellinger, Guy Cheron, and Thierry Dutoit. 2012. “A Subjective Assessment of a P300 BCI System for Lower-Limb Rehabilitation Purposes.” In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* 2012:3845–49. doi:10.1109/EMBC.2012.6346806.

Ehlis, A-C, M J Herrmann, A Wagener, and a J Fallgatter. 2005. “Multi-Channel near-Infrared Spectroscopy Detects Specific Inferior-Frontal Activation during Incongruent Stroop Trials.” *Biological Psychology* 69 (3): 315–31. doi:10.1016/j.biopsycho.2004.09.003.

Etkin, Amit, Tobias Egner, and Raffael Kalisch. 2011. “Emotional Processing in Anterior Cingulate and Medial Prefrontal Cortex.” *Trends in Cognitive Sciences* 15 (2). Elsevier Ltd: 85–93. doi:10.1016/j.tics.2010.11.004.

Falk, TH, M. Guirgis, S. Power, and TT Chau. 2011. “Taking NIRS-BCIs Outside the Lab: Towards Achieving Robustness against Environment Noise.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 19 (2): 136–46. doi:10.1109/TNSRE.2010.2078516.

Falk, Tiago H., Julie Chan, Pierre Duez, Gail Teachman, and Tom Chau. 2010. “Augmentative Communication Based on Realtime Vocal Cord Vibration Detection.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18 (2): 159–63. doi:10.1109/TNSRE.2009.2039593.

Faress, Ahmed, and Tom Chau. 2013. “Towards a Multimodal Brain-Computer Interface: Combining fNIRS and fTCD Measurements to Enable Higher Classification Accuracy.” *NeuroImage* 77. Elsevier Inc.: 186–94. doi:10.1016/j.neuroimage.2013.03.028.

Felton, E.A Williams, J.C Vanderheiden, G.C Radwin, R.G. 2012. “Mental Workload during Brain-Computer Interface Training.” *Ergonomics* 55 (5): 526–37. doi:10.1080/00140139.2012.662526.Mental.

Fenton, A., and S. Alpert. 2008. “Extending Our View on Using BCIs for Locked-in Syndrome.” *Neuroethics* 1: 119–32.

Ferrari, Marco, Leonardo Mottola, and Valentina Quaresima. 2004. “Principles, Techniques, and Limitations of near Infrared Spectroscopy.” *Canadian Journal of Applied Physiology* 29 (4): 463–87. http://www.nrcresearchpress.com/doi/abs/10.1139/h04-031.

Fitts, P M, and M I Posner. 1967. *Human Performance*. Brooks/Cole Publishing Company.

Friedrich, E V C, R Scherer, and C Neuper. 2013. “User-Appropriate and Robust Control Strategies to Enhance Brain−Computer Interface Performance and Usability.” In *Brain-Computer Interface Research SpringerBriefs in Electrical and Computer Engineering*, edited by Christoph Guger, Brendan Z. Allison, and Günter Edlinger, 15–22. SpringerBriefs in Electrical and Computer Engineering. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-36083-1.

Friedrich, Elisabeth V C, Reinhold Scherer, and Christa Neuper. 2012. “The Effect of Distinct Mental Strategies on Classification Performance for Brain-Computer Interfaces.” *International Journal of Psychophysiology* 84 (1). Elsevier B.V.: 86–94. doi:10.1016/j.ijpsycho.2012.01.014.

Friedrich, Elisabeth V. C., Christa Neuper, and Reinhold Scherer. 2013. “Whatever Works: A Systematic User-Centered Training Protocol to Optimize Brain-Computer Interfacing Individually.” Edited by Natasha M. Maurits. *PLoS ONE* 8 (9): e76214. doi:10.1371/journal.pone.0076214.

Funane, Tsukasa, Hirokazu Atsumori, Takusige Katura, Akiko N Obata, Hiroki Sato, Yukari Tanikawa, Eiji Okada, and Masashi Kiguchi. 2014. “Quantitative Evaluation of Deep and Shallow Tissue Layers’ Contribution to fNIRS Signal Using Multi-Distance Optodes and Independent Component Analysis.” *NeuroImage* 85. Elsevier Inc.: 150–65. doi:10.1016/j.neuroimage.2013.02.026.

Gallegos-Ayala, Guillermo, Adrian Furdea, Kouji Takano, Carolin A. Ruf, Herta Flor, and Niels Birbaumer. 2014. “Brain Communication in a Completely Locked-in Patient Using Bedside near-Infrared Spectroscopy.” *Neurology* 82 (21): 1930–32.

Gao, Wen-jun, Huai-xing Wang, Melissa A Snyder, and Yan-chun Li. 1990. “The Unique Properties of the Prefrontal Cortex and Mental Illness.” In *When Things Go Wrong - Diseases and Disorders of the Human Brain*, 3–26. InTech.

Gobel, Eric W, Todd B Parrish, and Paul J Reber. 2011. “Neural Correlates of Skill Acquisition: Decreased Cortical Activity during a Serial Interception Sequence Learning Task.” *NeuroImage* 58 (4). Elsevier Inc.: 1150–57. doi:10.1016/j.neuroimage.2011.06.090.

Graham, Steven, Jiaying Jiang, Victoria Manning, Ayna Baladi Nejad, Koh Zhisheng, Shan R Salleh, Xavier Golay, Yeh Ing Berne, and Peter J McKenna. 2010. “IQ-Related fMRI Differences during Cognitive Set Shifting.” *Cerebral Cortex* 20 (3): 641–49. doi:10.1093/cercor/bhp130.

Grosse-wentrup, Moritz, and Bernhard Schölkopf. 2013. “A Review of Performance Variations in SMR-Based Brain−Computer Interfaces (BCIs).” In *Brain-Computer Interface Research: A State-of-the-Art Summary*, edited by Christoph Guger, Brendan Z. Allison, and Günter Edlinger, 39–51. SpringerBriefs in Electrical and Computer Engineering. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-36083-1.

Guger, Christoph, Shahab Daban, Eric Sellers, Clemens Holzner, Gunther Krausz, Roberta Carabalona, Furio Gramatica, and Guenter Edlinger. 2009. “How Many People Are Able to Control a P300-Based Brain-Computer Interface (BCI)?” *Neuroscience Letters* 462 (1): 94–98. doi:10.1016/j.neulet.2009.06.045.

Gupta, A, S. Parameswaran, and Cheng-Han Lee. 2009. “Classification of Electroencephalography (EEG) Signals for Different Mental Activities Using Kullback Leibler (KL) Divergence.” *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1697–1700. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4959929.

Haeussinger, Florian B, Sebastian Heinzel, Tim Hahn, Martin Schecklmann, Ann-Christine Ehlis, and Andreas J Fallgatter. 2011. “Simulation of near-Infrared Light Absorption Considering Individual Head and Prefrontal Cortex Anatomy: Implications for Optical Neuroimaging.” *PloS One* 6 (10): e26377. doi:10.1371/journal.pone.0026377.

Haier, RJ, BENJAMIN Siegel, CHUCK Tang, LENNART Abel, and MS Buchsbaum. 1992. “Intelligence and Changes in Regional Cerebral Glucose Metabolic Rate Following

Learning." *Intelligence* 16: 415–26. http://www.sciencedirect.com/science/article/pii/016028969290018M.

Hammer, Eva Maria, Sebastian Halder, Benjamin Blankertzc, Claudia Sannelli, Thorsten Dickhaus, Sonja Kleihb, Klaus-Robert Müller, and Andrea Kübler. 2012. "Psychological Predictors of SMR-BCI Performance." *Biological Psychology* 89 (1). Elsevier B.V.: 80–86. doi:10.1016/j.biopsycho.2011.09.006.

Hart, S. G. 2006. "NASA-Task Load Index (NASA-TLX); 20 Years Later." In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, 904–8. http://humansystems.arc.nasa.gov/groups/tlx/.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. "Overview of Supervised Learning." In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 22–28. Springer, New York.

Hemphill, James F. 2003. "Interpreting the Magnitudes of Correlation Coefficients." *American Psychologist* 58 (1): 78–80. doi:10.1037/0003-066X.58.1.78.

Herff, Christian, Dominic Heger, Ole Fortmann, Johannes Hennrich, Felix Putze, and Tanja Schultz. 2013. "Mental Workload during N-Back Task-Quantified in the Prefrontal Cortex Using fNIRS." *Frontiers in Human Neuroscience* 7: 935. doi:10.3389/fnhum.2013.00935.

Herff, Christian, Dominic Heger, Felix Putze, Johannes Hennrich, Ole Fortmann, and Tanja Schultz. 2013. "Classification of Mental Tasks in the Prefrontal Cortex Using fNIRS." *Conference Proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, 2160–63. doi:10.1109/EMBC.2013.6609962.

Herrmann, M.J., A.-C. Ehlis, and A.J. Fallgatter. 2003. "Frontal Activation during a Verbal-Fluency Task as Measured by near-Infrared Spectroscopy." *Brain Research Bulletin* 61 (1): 51–56. doi:10.1016/S0361-9230(03)00066-2.

Hirshfield, LM Leanne M., Erin Treacy ET Solovey, Audrey Girouard, James Kebinger, Robert J.K. Jacob, Angelo Sassaroli, and Sergio Fantini. 2009. "Brain Measurement for Usability

Testing and Adaptive Interfaces : An Example of Uncovering Syntactic Workload with Functional Near Infrared Spectroscopy.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2185–94. http://dl.acm.org/citation.cfm?id=1519035.

Holz, Elisa Mira, Johannes Höhne, Pit Staiger-Sälzer, Michael Tangermann, and Andrea Kübler. 2013. “Brain-Computer Interface Controlled Gaming: Evaluation of Usability by Severely Motor Restricted End-Users.” *Artificial Intelligence in Medicine* 59 (2). Elsevier B.V.: 111–20. doi:10.1016/j.artmed.2013.08.001.

Hoonakker, Peter, Pascale Carayon, Ayse Gurses, Roger Brown, Adjhaporn Khunlertkit, Kerry McGuire, and James M Walker. 2011. “Measuring Workload of Icu Nurses With a Questionnaire Survey: The Nasa Task Load Index (Tlx).” *IIE Transactions on Healthcare Systems Engineering* 1 (2): 131–43. doi:10.1080/19488300.2011.609524.

Hoshi, Y, O Hirotaka, Y Watanabe, J Andersson, M Bergstroem, A Lilja, B Laenstroem, and M Tamura. 1994. “Non-Synchronous Behavior of Neuronal Activity, Oxidative Metabolism and Blood Supply during Mental Tasks in Man.” *Neuroscience Letters* 172: 129–33.

Hoshi, Yoko, Jinghua Huang, Shunji Kohri, Yoshinobu Iguchi, Masayuki Naya, Takahiro Okamoto, and Shuji Ono. 2011. “Recognition of Human Emotions from Cerebral Blood Flow Changes in the Frontal Region: A Study with Event-Related near-Infrared Spectroscopy.” *Journal of Neuroimaging* 21 (2): 94–101. doi:10.1111/j.1552-6569.2009.00454.x.

Husband, SD, and CM DeCato. 1982. “The Quick Test Compared with the Wechsler Adult Intelligence Scale as Measures of Intellectual Functioning in a Prison Clinical Setting.” *Psychological Reports* 40 (1): 167–70.

ISO 9241-11. 1998. *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) -- Part 11: Guidance on Usability*.

ISS Inc. 2012. “ISS Imagent: Functional Brain Imaging System Using Infrared Photons.” http://iss.com/biomedical/instruments/imagent.html.

Izzetoglu, Kurtulus, Hasan Ayaz, Anna Merzagora, Meltem Izzetoglu, Patricia a. Shewokis, Scott C. Bunce, Kambiz Pourrezaei, Arye Rosen, and Banu Onaral. 2011. “The Evolution Of Field Deployable fNIR Spectroscopy From Bench To Clinical Settings.” *Journal of Innovative Optical Health Sciences* 04 (03): 239–50. doi:10.1142/S1793545811001587.

Izzetoglu, M, K. Izzetoglu, S. Bunce, H. Ayaz, A. Devaraj, B. Onaral, and Kambiz Pourrezaei. 2005. “Functional near-Infrared Neuroimaging.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 13 (2): 153–59. doi:10.1109/TNSRE.2005.847377.

Izzetoglu, Meltem, Scott C Bunce, Kurtulus Izzetoglu, Banu Onaral, and Kambiz Pourrezaei. 2007. “Functional Brain Imaging Using near-Infrared Technology.” *IEEE Engineering in Medicine and Biology Magazine* 26 (4): 38–46. http://www.ncbi.nlm.nih.gov/pubmed/17672230.

Jain, a., and D. Zongker. 1997. “Feature Selection: Evaluation, Application, and Small Sample Performance.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (2): 153–58. doi:10.1109/34.574797.

Kaltenbach, Hans-Michael. 2012. “Hypothesis Testing Multiple Testing: Bonferroni-Correction.” In *A Concise Guide to Statistics*, 71–75. Springer. doi:10.1007/978-3-642-23502-3.

Kanoh, Shin’ichiro, Yu-Mi Murayama, Ko-Ichiro Miyamoto, Tatsuo Yoshinobu, and Ryuta Kawashima. 2009. “A NIRS-Based Brain-Computer Interface System during Motor Imagery: System Development and Online Feedback Training.” In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 594–97. doi:10.1109/IEMBS.2009.5333710.

Kanthack, Thiago F Dias, M Bigliassi, and LR Altimari. 2013. “Equal Prefrontal Cortex Activation between Males and Females in a Motor Tasks and Different Visual Imagery

Perspectives: A Functional near- Infrared Spectroscopy (fNIRS) Study." *Motriz Rio Claro* 19 (3): 627–32.

Khan, M Jawad, Melissa Jiyoun Hong, and Keum-Shik Hong. 2014. "Decoding of Four Movement Directions Using Hybrid NIRS-EEG Brain-Computer Interface." *Frontiers in Human Neuroscience* 8: Article 244. doi:10.3389/fnhum.2014.00244.

Kirkpatrick, P J, P Smielewski, P C Whitfield, M Czosnyka, D Menon, and J D Pickard. 1995. "An Observational Study of near-Infrared Spectroscopy during Carotid Endarterectomy." *Journal of Neurosurgery* 82 (5): 756–63. doi:10.3171/jns.1995.82.5.0756.

Kleih, S C, A Riccio, D Mattia, V Kaiser, E V C Friedrich, R Scherer, Gernot Mueller-Putz, C. Neuper, and A. Kuebler. 2011. "Motivation Influences Performance in SMR-BCI." In *Proceedings of the 5th International Brain-Computer Interface Conference*, 05:108–11.

Koch, Stefan P, Sophie Koendgen, Riad Bourayou, Jens Steinbrink, and Hellmuth Obrig. 2008. "Individual Alpha-Frequency Correlates with Amplitude of Visual Evoked Potential and Hemodynamic Response." *NeuroImage* 41 (2): 233–42.

Koegel, R L, and M Mentis. 1985. "Motivation in Childhood Autism: Can They or Won't They?" *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 26 (2): 185–91.

Koelstra, Sander, Ashkan Yazdani, Mohammad Soleymani, Christian Mühl, Jong-Seok Lee, Anton Nijholt, Thierry Pun, Touradj Ebrahimi, and Ioannis Patras. 2010. "Single Trial Classification of EEG and Peripheral Physiological Signals for Recognition of Emotions Induced by Music Videos." *Brain Informatics Lecture Notes in Computer Science* 6334: 89–100. http://link.springer.com/chapter/10.1007/978-3-642-15314-3_9.

Kopton, Isabella M, and Peter Kenning. 2014. "Near-Infrared Spectroscopy (NIRS) as a New Tool for Neuroeconomic Research." *Frontiers in Human Neuroscience* 8: Article 549. doi:10.3389/fnhum.2014.00549.

Koshino, Hideya, Takehiro Minamoto, Takashi Ikeda, M Osaka, Y Otsuka, and N Osaka. 2011. "Anterior Medial Prefrontal Cortex Exhibits Activation during Task Preparation but

Deactivation during Task Execution." *PloS One* 6 (8): e22909. doi:10.1371/journal.pone.0022909.

Kotchoubey, B., D. Schneider, H. Schleichert, U. Strehl, C. Uhlmann, W. Blankenhorn, W. Froescher, and N. Birbaumer. 1996. "Self-Regulation of Slow Cortical Potentials in Epilepsy: A Retrial with Analysis of Influencing Factors." *Epilepsy Research* 25 (3): 269–76. http://resolver.scholarsportal.info/resolve/09201211/v25i0003/269_soscpirwaoif.xml.

Kübler, A, B Kotchoubey, T Hinterberger, N Ghanayim, J Perelmouter, M Schauer, C Fritsch, E Taub, and N Birbaumer. 1999. "The Thought Translation Device: A Neurophysiological Approach to Communication in Total Motor Paralysis." *Experimental Brain Research* 124 (2): 223–32. http://www.ncbi.nlm.nih.gov/pubmed/9928845.

Kübler, A, F Nijboer, J Mellinger, T M Vaughan, H Pawelzik, G Schalk, D J Mcfarland, N Birbaumer, and J R Wolpaw. 2005. "Patients with ALS Can Use Sensorimotor Rhythms to Operate a Brain-." *Neurology* 64 (10): 1775–77.

Kübler, A., and N. Birbaumer. 2008. "Brain-Computer Interfaces and Communication in Paralysis: Extinction of Goal Directed Thinking in Completely Paralysed Patients?" *Health Psychology* 119 (11): 2658–66. doi:10.1016/j.clinph.2008.06.019.Brain-Computer.

Kübler, Andrea, Adrian Furdea, Sebastian Halder, Eva Maria Hammer, Femke Nijboer, and Boris Kotchoubey. 2009. "A Brain-Computer Interface Controlled Auditory Event-Related Potential (p300) Spelling System for Locked-in Patients." *Annals of the New York Academy of Sciences* 1157: 90–100. doi:10.1111/j.1749-6632.2008.04122.x.

Kübler, Andrea, Boris Kotchoubey, Jochen Kaiser, Jonathan R. Wolpaw, and Niels Birbaumer. 2001. "Brain–computer Communication: Unlocking the Locked in." *Psychological Bulletin* 127 (3): 358–75. http://resolver.scholarsportal.info/resolve/00332909/v127i0003/358_bcutli.xml.

Kübler, Andrea, Nicola Neumann, Jochen Kaiser, Boris Kotchoubey, Thilo Hinterberger, and Niels P. Birbaumer. 2001. "Brain-Computer Communication: Self-Regulation of Slow

Cortical Potentials for Verbal Communication.” *Archives of Physical Medicine and Rehabilitation* 82 (11): 1533–39. doi:10.1053/apmr.2001.26621.

Kudo, Mineichi, and Jack Sklansky. 2000. “Comparison of Algorithms That Select Features for Pattern Classifiers.” *Pattern Recognition* 33 (1): 25–41. doi:10.1016/S0031-3203(99)00041-2.

Kwee, IL, and Tsutomu Nakada. 2003. “Dorsolateral Prefrontal Lobe Activation Declines Significantly with Age--Functional NIRS Study.” *Journal of Neurology* 250 (5): 525–29. doi:10.1007/s00415-003-1028-x.

Langan-fox, Janice, Kathy Armstrong, Nikola Balvin, and Jeromy Anglim. 2002. “Process in Skill Acquisition: Motivation, Interruptions, Memory, Affective States, and Metacognition.” *Australian Pyschologist* 37 (2): 104–17.

Leff, Daniel Richard, Felipe Orihuela-Espina, Clare E Elwell, Thanos Athanasiou, David T Delpy, Ara W Darzi, and Guang-Zhong Yang. 2011. “Assessment of the Cerebral Cortex during Motor Task Behaviours in Adults: A Systematic Review of Functional near Infrared Spectroscopy (fNIRS) Studies.” *NeuroImage* 54 (4). Elsevier Inc.: 2922–36. doi:10.1016/j.neuroimage.2010.10.058.

Leung, Brian, Jessica A Brian, and Tom Chau. 2013. “Learning and Mastery Behaviours as Risk Factors to Abandonment in a Paediatric User of Advanced Single-Switch Access Technology.” *Disability and Rehabilitation. Assistive Technology* 8 (5): 426–33. doi:10.3109/17483107.2012.754955.

Lillo, Patricia, and John R Hodges. 2010. “Cognition and Behaviour in Motor Neurone Disease.” *Current Opinion in Neurology* 23 (6): 638–42. doi:10.1097/WCO.0b013e3283400b41.

Lloyd-Fox, S, A Blasi, and C E Elwell. 2010. “Illuminating the Developing Brain: The Past, Present and Future of Functional near Infrared Spectroscopy.” *Neuroscience and Biobehavioral Reviews* 34 (3): 269–84. doi:10.1016/j.neubiorev.2009.07.008.

Mak, J N, and J R Wolpaw. 2009. “Clinical Applications of Brain-Computer Interfaces: Current State and Future Prospects.” *IEEE Reviews in Biomedical Engineering* 2: 187–99. doi:10.1109/RBME.2009.2035356.Clinical.

Marjoram, Dominic, Clare Gardner, Jonathan Burns, Patrick Miller, Stephen M Lawrie, and Eve C Johnstone. 2005. “Symptomatology and Social Inference: A Theory of Mind Study of Schizophrenia and Psychotic Affective Disorder.” *Cognitive Neuropsychiatry* 10 (5): 347–59. doi:10.1080/13546800444000092.

Marumo, Kohei, Ryu Takizawa, Yuki Kawakubo, Toshiaki Onitsuka, and Kiyoto Kasai. 2009. “Gender Difference in Right Lateral Prefrontal Hemodynamic Response While Viewing Fearful Faces: A Multi-Channel near-Infrared Spectroscopy Study.” *Neuroscience Research* 63 (2): 89–94. doi:10.1016/j.neures.2008.10.012.

Mason, S G, and G E Birch. 2000. “A Brain-Controlled Switch for Asynchronous Control Applications.” *IEEE Transactions on Bio-Medical Engineering* 47 (10): 1297–1307. doi:10.1109/10.871402.

McColl, M. a., and H. Skinner. 1995. “Assessing Inter- and Intrapersonal Resources: Social Support and Coping among Adults with a Disability.” *Disability & Rehabilitation* 17 (1): 24–34. doi:10.3109/09638289509166624.

Moghimi, Saba, Azadeh Kushki, Anne Marie Guerguerian, and Tom Chau. 2012. “A Review of EEG-Based Brain-Computer Interfaces as Access Pathways for Individuals with Severe Disabilities.” *Assistive Technology*. doi:10.1080/10400435.2012.723298.

Moghimi, Saba, Azadeh Kushki, Sarah Power, Anne Marie Guerguerian, and Tom Chau. 2012. “Automatic Detection of Prefrontal Cortical Response to Emotionally Rated Music Using Multi-Channel near-Infrared Spectroscopy.” *Journal of Neural Engineering* 9 (2): 026022.

Moosmann, Matthias, Petra Ritter, Ina Krastel, Andrea Brink, Sebastian Thees, Felix Blankenburg, Birol Taskin, Hellmuth Obrig, and Arno Villringer. 2003. “Correlates of

Alpha Rhythm in Functional Magnetic Resonance Imaging and near Infrared Spectroscopy.” *NeuroImage* 20 (1): 145–58. doi:10.1016/S1053-8119(03)00344-6.

Morshed, Bashir I, and Abdulhalim Khan. 2014. “Biomedical Science A Brief Review of Brain Signal Monitoring Technologies for BCI Applications : Challenges and Prospects” 4 (1): 1–10. doi:10.4172/2155-9538.1000141.

Mueller-Putz, Gernot, Reinhold Scherer, Clemens Brunner, Robert Leeb, and Gert Pfurtscheller. 2008. “Better than Random? A Closer Look on BCI Results.” *International Journal of Bioelectromagnetism* 10 (1): 52–55. http://ijbem.ws.hosei.ac.jp/volume10/number1/100107.pdf.

Nai-Jen, Huan, and R. Palaniappan. 2004. “Classification of Mental Tasks Using Fixed and Adaptive Autoregressive Models of EEG Signals.” In *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, 1:507–10. doi:10.1109/IEMBS.2004.1403205.

Naito, Masayoshi, Yohko Michioka, Kuniaki Ozawa, Yoshitoshi Ito, Masashi Kiguchi, and Tsuneo Kanazawa. 2007. “A Communication Means for Totally Locked-in ALS Patients Based on Changes in Cerebral Blood Volume Measured with Near-Infrared Light.” *IEICE - Transactions on Information and Systems* E90-D (7): 1028–37. http://search.ieice.org/bin/summary.php?id=e90-d_7_1028.

Naseer, Noman, and Keum-Shik Hong. 2013a. “Functional near-Infrared Spectroscopy Based Discrimination of Mental Counting and No-Control State for Development of a Brain-Computer Interface.” In *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2013:1780–83. doi:10.1109/EMBC.2013.6609866.

Naseer, Noman, and Keum-Shik Hong. 2013b. “Classification of Functional near-Infrared Spectroscopy Signals Corresponding to the Right- and Left-Wrist Motor Imagery for Development of a Brain-Computer Interface.” *Neuroscience Letters* 553. Elsevier Ireland Ltd: 84–89. doi:10.1016/j.neulet.2013.08.021.

Naseer, Noman, Melissa Jiyoun Hong, and Keum-Shik Hong. 2014. "Online Binary Decision Decoding Using Functional near-Infrared Spectroscopy for the Development of Brain-Computer Interface." *Experimental Brain Research* 232 (2): 555–64. doi:10.1007/s00221-013-3764-1.

Neary, D, J.S Snowden, and D.M.a Mann. 2000. "Cognitive Change in Motor Neurone Disease/amyotrophic Lateral Sclerosis (MND/ALS)." *Journal of the Neurological Sciences* 180 (1-2): 15–20. doi:10.1016/S0022-510X(00)00425-1.

Newman, Sharlene D, Timothy a Keller, and Marcel Adam Just. 2007. "Volitional Control of Attention and Brain Activation in Dual Task Performance." *Human Brain Mapping* 28 (2): 109–17. doi:10.1002/hbm.20257.

Nicolas-Alonso, Luis Fernando, and Jaime Gomez-Gil. 2012. "Brain Computer Interfaces, a Review." *Sensors* 12 (2): 1211–79. doi:10.3390/s120201211.

Nie, Jingxin, Lei Guo, and Tianming Liu. 2009. "A Computational Model of Cerebral Cortex Folding." *Medical Image Computing and Computer-Assisted Intervention* 12 (2): 458–65. http://www.sciencedirect.com/science/article/pii/S0022519310000615.

Nijboer, Femke, Niels Birbaumer, and Andrea Kübler. 2010. "The Influence of Psychological State and Motivation on Brain-Computer Interface Performance in Patients with Amyotrophic Lateral Sclerosis - a Longitudinal Study." *Frontiers in Neuroscience* 4 (Article 55): 1–13. doi:10.3389/fnins.2010.00055.

Nijboer, Femke, Adrian Furdea, Ingo Gunst, Jürgen Mellinger, Dennis J McFarland, Niels Birbaumer, and Andrea Kübler. 2008. "An Auditory Brain-Computer Interface (BCI)." *Journal of Neuroscience Methods* 167 (1): 43–50. doi:10.1016/j.jneumeth.2007.02.009.

O'Reilly, Randall C. 2010. "The What and How of Prefrontal Cortical Organization." *Trends in Neurosciences* 33 (8). Elsevier Ltd: 355–61. doi:10.1016/j.tins.2010.05.002.

Oberg, Ann L., and Douglas W. Mahoney. 2007. "Linear Mixed Effects Models." In *Topics in Biostatistics*, edited by Walter T Ambrosius, 213–34. Humana Press.

Ogata, H., T. Mukai, and T. Yagi. 2007. "A Study on the Frontal Cortex in Cognitive Tasks Using Near-Infrared Spectroscopy." In *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, 4731–34. doi:10.1109/IEMBS.2007.4353396.

Okada, E, M Firbank, M Schweiger, S R Arridge, M Cope, and D T Delpy. 1997. "Theoretical and Experimental Investigation of near-Infrared Light Propagation in a Model of the Adult Head." *Applied Optics* 36 (1): 21–31. http://www.ncbi.nlm.nih.gov/pubmed/18250644.

Okada, Fumihiko, Yukiko Tokumitsu, Yoko Hoshi, and Mamoru Tamura. 1993. "Gender- and Handedness-Related Differences of Forebrain Oxygenation and Hemodynamics." *Brain Research* 601 (1–2): 337–42. doi:10.1016/0006-8993(93)91733-9.

Oldfield, R C. 1971. "The Assessment and Analysis of Handedness: The Edinburgh Inventory." *Neuropsychologia* 9 (1): 97–113. http://www.ncbi.nlm.nih.gov/pubmed/5146491.

Palaniappan, Ramaswamy. 2006. "Utilizing Gamma Band to Improve Mental Task Based Brain-Computer Interface Design." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 14 (3): 299–303. doi:10.1109/TNSRE.2006.881539.

Palaniappan, Ramaswamy, Raveendran Paramesran, Shogo Nishida, and Naoki Saiwaki. 2002. "A New Brain-Computer Interface Design Using Fuzzy ARTMAP." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 10 (3): 140–48. doi:10.1109/TNSRE.2002.802854.

Pfurtscheller, G, C Neuper, G R Müller, B Obermaier, G Krausz, A Schlögl, R Scherer, et al. 2003. "Graz-BCI: State of the Art and Clinical Applications." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 11 (2): 177–80. doi:10.1109/TNSRE.2003.814454.

Pfurtscheller, Gert, Brendan Z Allison, Clemens Brunner, Gunther Bauernfeind, Teodoro Solis-Escalante, Reinhold Scherer, Thorsten O Zander, Gernot Mueller-Putz, Christa Neuper, and Niels Birbaumer. 2010. "The Hybrid BCI." *Frontiers in Neuroscience* 4: Article 30. doi:10.3389/fnpro.2010.00003.

Pfurtscheller, Gert, Günther Bauernfeind, Selina Christin Wriessnegger, and Christa Neuper. 2010. "Focal Frontal (de)oxyhemoglobin Responses during Simple Arithmetic." *International Journal of Psychophysiology* 76 (3). Elsevier B.V.: 186–92. doi:10.1016/j.ijpsycho.2010.03.013.

Pineda, Jaime A, David S Silverman, Andrey Vankov, and John Hestenes. 2003. "Learning to Control Brain Rhythms: Making a Brain-Computer Interface Possible." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 11 (2): 181–84. doi:10.1109/TNSRE.2003.814445.

Poldrack, Russell a, Fred W Sabb, Karin Foerde, Sabrina M Tom, Robert F Asarnow, Susan Y Bookheimer, and Barbara J Knowlton. 2005. "The Neural Correlates of Motor Skill Automaticity." *The Journal of Neuroscience* 25 (22): 5356–64. doi:10.1523/JNEUROSCI.3880-04.2005.

Polikar, R. 2006. "Ensemble Based Systems in Decision Making." *Circuits and Systems Magazine, IEEE* 6: 21–45. doi:10.1109/MCAS.2006.1688199.

Power, Sarah D, Azadeh Kushki, and Tom Chau. 2012. "Automatic Single-Trial Discrimination of Mental Arithmetic, Mental Singing and the No-Control State from Prefrontal Activity: Toward a Three-State NIRS-BCI." *BMC Research Notes* 5 (141): 1–10. doi:10.1186/1756-0500-5-141.

Power, Sarah D., Tiago H. Falk, and Tom Chau. 2010. "Classification of Prefrontal Activity due to Mental Arithmetic and Music Imagery Using Hidden Markov Models and Frequency Domain near-Infrared Spectroscopy." *Journal of Neural Engineering* 7 (2): 026002. doi:10.1088/1741-2560/7/2/026002.

Power, Sarah D., Azadeh Kushki, and Tom Chau. 2012. "Intersession Consistency of Single-Trial Classification of the Prefrontal Response to Mental Arithmetic and the No-Control State by NIRS." *PloS One* 7 (7): e37791. doi:10.1371/journal.pone.0037791.

Power, Sarah Dianne, and Tom Chau. 2013. “Automatic Single-Trial Classification of Prefrontal Hemodynamic Activity in an Individual with Duchenne Muscular Dystrophy.” *Developmental Neurorehabilitation* 16 (1): 67–72. doi:10.3109/17518423.2012.718293.

Power, Sarah, Azadeh Kushki, and Tom Chau. 2011. “Towards a System-Paced near-Infrared Spectroscopy Brain-Computer Interface: Differentiating Prefrontal Activity due to Mental Arithmetic and Mental Singing from the No-Control State.” *Journal of Neural Engineering* 8 (6): 066004. doi:10.1088/1741-2560/8/6/066004.

Pudil, P., J. Novovičová, and J. Kittler. 1994. “Floating Search Methods in Feature Selection.” *Pattern Recognition Letters* 15 (11). Elsevier Science Inc.: 1119–25. doi:10.1016/0167-8655(94)90127-9.

Quaresima, Valentina, Marco Ferrari, Alessandro Torricelli, Lorenzo Spinelli, Antonio Pifferi, and Rinaldo Cubeddu. 2005. “Bilateral Prefrontal Cortex Oxygenation Responses to a Verbal Fluency Task: A Multichannel Time-Resolved near-Infrared Topography Study.” *Journal of Biomedical Optics* 10 (1): 11012. doi:10.1117/1.1851512.

Quijano-Roy, S, I Martí-Carrera, S Makri, M Mayer, S Maugenre, P Richard, C Berard, et al. 2006. “Brain MRI Abnormalities in Muscular Dystrophy due to FKRP Mutations.” *Brain and Development* 28: 232–42. http://www.sciencedirect.com/science/article/pii/S0387760405001774.

Quinlan, Philip, and Ben Dyson. 2008. “Chapter 8 Attention: General Introduction, Basic Models and Data.” In *Cognitive Psychology*, 271–311. Person Education Limited, England.

Rahimi-Golkhandan, S., J.P. Piek, B. Steenbergen, and P.H. Wilson. 2014. “Hot Executive Function in Children with Developmental Coordination Disorder: Evidence for Heightened Sensitivity to Immediate Reward.” *Cognitive Development* 32. Elsevier Inc.: 23–37. doi:10.1016/j.cogdev.2014.06.002.

Rahm, Christoffer, Benny Liberg, Maria Wiberg-Kristoffersen, Peter Aspelin, and Mussie Msghina. 2013. “Rostro-Caudal and Dorso-Ventral Gradients in Medial and Lateral

Prefrontal Cortex during Cognitive Control of Affective and Cognitive Interference." *Scandinavian Journal of Psychology* 54 (2): 66–71. doi:10.1111/sjop.12023.

Randolph, Adriane, Saurav Karmakar, and Melody Jackson. 2006. "Towards Predicting Control of a Brain-Computer Interface." In *ICIS 2006 Proceedings*, Paper 53.

Refaeilzadeh, Payam, Lei Tang, and Huan Liu. 2009. "Cross-Validation." *Encyclopedia of Database Systems*. Springer US. doi:10.1159/000323569.

Roca, Juan Carlos, Chao-Min Chiu, and Francisco José Martínez. 2006. "Understanding E-Learning Continuance Intention: An Extension of the Technology Acceptance Model." *International Journal of Human-Computer Studies* 64 (8): 683–96. doi:10.1016/j.ijhcs.2006.01.003.

Rocha, Anderson, and SK Goldenstein. 2014. "Multiclass from Binary: Expanding One-Versus-All, One-Versus-One and ECOC-Based Approaches." *IEEE Transactions on Neural Networks and Learning Systems* 25 (2): 289–302. doi:10.1109/TNNLS.2013.2274735.

Roche-Labarbe, Nadège, Boubker Zaaimi, Mahdi Mahmoudzadeh, Victoria Osharina, Alice Wallois, Astrid Nehlig, Reinhard Grebe, and Fabrice Wallois. 2010. "NIRS-Measured Oxy- and Deoxyhemoglobin Changes Associated with EEG Spike-and-Wave Discharges in a Genetic Model of Absence Epilepsy: The GAERS." *Epilepsia* 51 (8): 1374–84. doi:10.1111/j.1528-1167.2010.02574.x.

Rogers, R D, a M Owen, H C Middleton, E J Williams, J D Pickard, B J Sahakian, and T W Robbins. 1999. "Choosing between Small, Likely Rewards and Large, Unlikely Rewards Activates Inferior and Orbital Prefrontal Cortex." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 19 (20): 9029–38. doi:0270-6474/99/19209029-10$05.00/0.

Romano, Jennifer C, James H Howard, and Darlene V Howard. 2010. "One-Year Retention of General and Sequence-Specific Skills in a Probabilistic, Serial Reaction Time Task." *Memory* 18 (4): 427–41. doi:10.1080/09658211003742680.

Rubio, Susana, Eva Díaz, Jesús Martín, and José M Puente. 2004. “Evaluation of Subjective Mental Workload : A Comparison of SWAT , NASA-TLX , and Workload Profile Methods.” *Applied Pyschology* 53 (1): 61–86.

Rypma, Bart, Jeffrey S Berger, and Mark D’Esposito. 2002. “The Influence of Working-Memory Demand and Subject Performance on Prefrontal Cortical Activity.” *Journal of Cognitive Neuroscience* 14 (5): 721–31. doi:10.1162/08989290260138627.

Rypma, BART, and MARK D’esposito. 1999. “The Roles of Prefrontal Brain Regions in Components of Working Memory : Effects of Memory Load and Individual Differences.” *Proceedings of the National Academy of Sciences of the United States of America PNAS* 96 (11): 6558–63.

Salden, Ron J.C.M., Fred Paas, and Jeroen J.G. van Merriënboer. 2006. “Personalised Adaptive Task Selection in Air Traffic Control: Effects on Training Efficiency and Transfer.” *Learning and Instruction* 16 (4): 350–62. doi:10.1016/j.learninstruc.2006.07.007.

Sasai, Shuntaro, Fumitaka Homae, Hama Watanabe, Akihiro T. Sasaki, Hiroki C. Tanabe, Norihiro Sadato, and Gentaro Taga. 2012. “A NIRS-fMRI Study of Resting State Network.” *NeuroImage* 63 (1): 179–93. doi:10.1016/j.neuroimage.2012.06.011.

Sato, H, N Yahata, T Funane, R Takizawa, T Katura, H Atsumori, Y Nishimura, et al. 2013. “A NIRS-fMRI Investigation of Prefrontal Cortex Activity during a Working Memory Task.” *NeuroImage* 83: 158–73. doi:10.1016/j.neuroimage.2013.06.043.

Sauro, Jeff, and Joseph S Dumas. 2009. “Comparison of Three One-Question , Post-Task Usability Questionnaires.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1599–1608.

Schlösser, R, M Hutchinson, S Joseffer, H Rusinek, A Saarimaki, J Stevenson, S L Dewey, and J D Brodie. 1998. “Functional Magnetic Resonance Imaging of Human Brain Activity in a Verbal Fluency Task.” *Journal of Neurology, Neurosurgery, and Psychiatry* 64 (4): 492–98.

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2170033&tool=pmcentrez&rendertype=abstract.

Scholkmann, Felix, Stefan Kleiser, Andreas Jaakko Metz, Raphael Zimmermann, Juan Mata Pavia, Ursula Wolf, and Martin Wolf. 2014. "A Review on Continuous Wave Functional near-Infrared Spectroscopy and Imaging Instrumentation and Methodology." *NeuroImage* 85 (1). Elsevier Inc.: 6–27. doi:10.1016/j.neuroimage.2013.05.004.

Schroeter, Matthias L, Stefan Zysset, Frithjof Kruggel, and D Yves von Cramon. 2003. "Age Dependency of the Hemodynamic Response as Measured by Functional near-Infrared Spectroscopy." *NeuroImage* 19 (3): 555–64. http://www.ncbi.nlm.nih.gov/pubmed/12880787.

Schroeter, Matthias L, Stefan Zysset, Thomas Kupka, Frithjof Kruggel, and D Yves von Cramon. 2002. "Near-Infrared Spectroscopy Can Detect Brain Activity during a Color-Word Matching Stroop Task in an Event-Related Design." *Human Brain Mapping* 17 (1): 61–71. doi:10.1002/hbm.10052.

Schudlo, Larissa C, and Tom Chau. 2014. "Dynamic Topographical Pattern Classification of Multichannel Prefrontal NIRS Signals : II . Online Differentiation of Mental Arithmetic and Rest." *Journal of Neural Engineering* 11 (1): 016003.

Schudlo, Larissa C, Sarah D Power, and Tom Chau. 2013. "Dynamic Topographical Pattern Classification of Multichannel Prefrontal NIRS Signals." *Journal of Neural Engineering* 10 (4): 046018. doi:10.1088/1741-2560/10/4/046018.

Schudlo, Larissa, Sabine Weyand, and Tom Chau. 2014. "A Review of Past and Future Near-Infrared Spectroscopy Brain Computer Interface Research at the PRISM Lab." In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1996–99.

Schulz, Richard, and Susan Decker. 1985. "Long-Term Adjustment to Physical Disability : The Role of Social Support , Perceived Control , and Self-Blame" 48 (1977): 1162–72.

Shapira-Lichter, Irit, Noga Oren, Yael Jacob, Michal Gruberger, and Talma Hendler. 2013. “Portraying the Unique Contribution of the Default Mode Network to Internally Driven Mnemonic Processes.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (13): 4950–55. doi:10.1073/pnas.1209888110.

Shiffrin, Richard M., and Walter Schneider. 1977. “Controlled and Automatic Human Information Processing: II Perceptual Learning, Automatic Attending, and a General Theory.” *Pyschological Review* 84 (2): 127–90.

Shin, Jaeyoung, and Jichai Jeong. 2014. “Multiclass Classification of Hemodynamic Responses for Performance Improvement of Functional near-Infrared Spectroscopy-Based Brain-Computer Interface.” *Journal of Biomedical Optics* 19 (6): 067009. doi:10.1117/1.JBO.19.6.067009.

Shin, Jaeyoung, Sungyong Kang, Minkyu Sung, Joohwan Kim, Yongjung Kim, Jihyun Kim, and Jichai Jeong. 2013. “A Study on Information Transfer Rate by Brain-Computer Interface (BCI) Using Functional near-Infrared Spectroscopy (fNIRS).” In *International Winter Workshop on Brain-Computer Interface (BCI)*, 32–33. Ieee. doi:10.1109/IWW-BCI.2013.6506620.

Sitaram, R., H. Zhang, C. Guan, M. Thulasidas, Yoko Hoshi, A. Ishikawa, K. Shimizu, and Niels Birbaumer. 2007. “Temporal Classification of Multichannel near-Infrared Spectroscopy Signals of Motor Imagery for Developing a Brain-Computer Interface.” *NeuroImage* 34 (4): 1416–27.

Skinner, B. F. 1948. “‘Superstition’ in the Pigeon.” *Journal of Experimental Psychology* 38 (2): 168–72. doi:10.1037/h0055873.

Smith, Anna B., Rozmin Halari, Vincent Giampetro, Michael Brammer, and Katya Rubia. 2011. “Developmental Effects of Reward on Sustained Attention Networks.” *NeuroImage* 56 (3). Elsevier Inc.: 1693–1704. doi:10.1016/j.neuroimage.2011.01.072.

Sorger, Bettina, Brigitte Dahmen, Joel Reithler, O Gosseries, A Maudoux, S Laureys, and R Goebel. 2009. “Another Kind of ‘BOLD Response’: Answering Multiple-Choice Questions via Online Decoded Single-Trial Brain Signals.” *Progress in Brain Research* 177: 275–92. doi:10.1016/S0079-6123(09)17719-1.

Stangl, Matthias, Günther Bauernfein, Jürgen Kurzmann, Reinhold Scherer, and Christa Neuper. 2013. “A Haemodynamic Brain-Computer Interface Based on Real-Time Classification of near Infrared Spectroscopy Signals during Motor Imagery and Mental Arithmetic.” *Journal of Near Infrared Spectroscopy* 21 (3): 157–71. doi:10.1255/jnirs.1048.

Strait, Megan, and Matthias Scheutz. 2014. “What We Can and Cannot (yet) Do with Functional near Infrared Spectroscopy.” *Frontiers in Neuroscience* 8: Article 117. doi:10.3389/fnins.2014.00117.

Sulzer, J, S Haller, F Scharnowski, N Weiskopf, N Birbaumer, M L Blefari, a B Bruehl, et al. 2013. “Real-Time fMRI Neurofeedback: Progress and Challenges.” *NeuroImage* 76. Elsevier Inc.: 386–99. doi:10.1016/j.neuroimage.2013.03.033.

Sun, Pei-Chen, Ray J. Tsai, Glenn Finger, Yueh-Yang Chen, and Dowming Yeh. 2008. “What Drives a Successful E-Learning? An Empirical Investigation of the Critical Factors Influencing Learner Satisfaction.” *Computers & Education* 50 (4): 1183–1202. doi:10.1016/j.compedu.2006.11.007.

Suzuki, Satoshi, Fumio Harashima, and Katsuhisa Furuta. 2010. “Human Control Law and Brain Activity of Voluntary Motion by Utilizing a Balancing Task with an Inverted Pendulum.” *Advances in Human-Computer Interaction* 2010 (Article ID 215825): 1–16. doi:10.1155/2010/215825.

Tachtsidis, Ilias, Terence S Leung, Laurence Devoto, David T Delpy, and Clare E Elwell. 2008. “Measurement of Frontal Lobe Functional Activation and Related Systemic Effects: A near-Infrared Spectroscopy Investigation.” *Advances in Experimental Medicine and Biology* 614: 397–403. doi:10.1007/978-0-387-74911-2_44.

Tai, K., S. Blain, and T. Chau. 2008. "A Review of Emerging Access Technologies for Individuals With Severe Motor Impairments." *Assistive Technology* 20 (4): 204–19.

Tai, Kelly, and Tom Chau. 2009. "Single-Trial Classification of NIRS Signals during Emotional Induction Tasks: Towards a Corporeal Machine Interface." *Journal of Neuroengineering and Rehabilitation* 6 (39): 1–14. doi:10.1186/1743-0003-6-39.

Takahashi, Toshimitsu, Yoriko Takikawa, Reiko Kawagoe, Satoshi Shibuya, Takayuki Iwano, and Shigeru Kitazawa. 2011. "Measurement of Frontal Lobe Functional Activation and Related Systemic Effects: A near-Infrared Spectroscopy Investigation." *Neuro Image* 57: 991–1002. doi:10.1016/j.neuroimage.2011.05.012.

Talukdar, M Tanveer, H Robert Frost, and Solomon G Diamond. 2015. "Modeling Neurovascular Coupling from Clustered Parameter Sets for Multimodal EEG-NIRS." *Computational and Mathematical Methods in Medicine* 2015 (830849): 12.

Tamura, M, Y Hoshi, and F Okada. 1997. "Localized near-Infrared Spectroscopy and Functional Optical Imaging of Brain Activity." *Philosophical Transactions of the Royal Society B: Biological Sciences* 352 (1354): 737–42. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1691956/.

Tan, D, and A Nijholt. 2010. *Brain-Computer Interfaces: Applying Our Minds to Human-Computer Interaction*. Springer.

Taylor, R. 1990. "Interpretation of the Correlation Coefficient: A Basic Review." *Journal of Diagnostic Medical Sonography* 6 (1): 35–39. doi:10.1177/875647939000600106.

Tedesco, Donna P, and Thomas S Tullis. 2006. "A Comparison of Methods for Eliciting Post-Task Subjective Ratings in Usability Testing." In *Usability Professionals Association Annual Conference (UPA)*, 1–9.

Thorndike, E. L., and R. T. Rock. 1934. "Learning without Awareness of What Is Being Learned or Intent to Learn It." *Journal of Experimental Psychology* 17 (1): 1–19. doi:10.1037/h0073815.

Toomim, H., W. Mize, P. Yeekwong, M. Toomim, R. Marsh, G. P. Kozlowski, M. Kimball, and A. Remond. 2005. "Intentional Increase of Cerebral Blood Oxygenation Using Hemoencephalography (HEG): An Efficient Brain Exercise Therapy." *Journal of Neurotherapy* 8 (3): 5–21.

Tripoliti, EE, FI Dimitrios, and M Argyropoulou. 2007. "A Supervised Method to Assist the Diagnosis of Alzheimer's Disease Based on Functional Magnetic Resonance Imaging." In *IEEE Engineering in Medicince and Biology Socity*, 3426–29.

Turkeltaub, Peter E, Lynn Gareau, D Lynn Flowers, Thomas a Zeffiro, and Guinevere F Eden. 2003. "Development of Neural Mechanisms for Reading." *Nature Neuroscience* 6 (7): 767–73. doi:10.1038/nn1065.

Utsugi, Kei, Akiko Obata, Hiroki Sato, Takusige Katsura, Kazuhiko Sagara, Atsushi Maki, and Hideaki Koizumi. 2007. "Development of an Optical Brain-Machine Interface." In *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, 5338–41. doi:10.1109/IEMBS.2007.4353547.

Vealey, RS, SW Hayashi, M Garner-Holman, and P Giacobbi. 1998. "Sources of Sport-Confidence: Conceptualization and Instrument Development." *Journal of Sport and Exercise Psychology* 20 (1): 54–80. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Sources+of+sport-confidence:+conceptualisation+and+instrument+development#0.

Villringer, A, J Planck, C Hock, L Schleinkofer, and U Dirnagl. 1993. "Near Infrared Spectroscopy (NIRS): A New Tool to Study Hemodynamic Changes during Activation of Brain Function in Human Adults." *Neuroscience Letters* 154 (1-2): 101–4. http://www.ncbi.nlm.nih.gov/pubmed/8361619.

Villringer, Arno, and Britton Chance. 1997. "Non-Invasive Optical Spectroscopy and Imaging of Human Brain Function." *Trends in Neurosciences* 20 (10): 435–42. doi:10.1016/S0166-2236(97)01132-6.

Wager, Tor D, and Edward E Smith. 2003. "Neuroimaging Studies of Working Memory: A Meta-Analysis." *Cognitive, Affective & Behavioral Neuroscience* 3 (4): 255–74. http://www.ncbi.nlm.nih.gov/pubmed/15040547.

Weierich, Mariann R., Christopher I. Wright, Alyson Negreira, Brad C. Dickerson, and Lisa Feldman Barrett. 2010. "Novelty as a Dimension in the Affective Brain." *NeuroImage* 49 (3). Elsevier Inc.: 2871–78. doi:10.1016/j.neuroimage.2009.09.047.

Weiskopf, N., R. Veit, M. Erb, K. Mathiak, W. Grodd, G. Rainer, and N. Birbaumer. 2003. "Physiological Self-Regulation of Regional Brain Activity Using Real-Time Functional Magnetic Resonance Imaging (fMRI): Methodology and Exemplary Data." *NeuroImage* 19 (3): 577–86.

Weiskopf, Nikolaus, Ranganatha Sitaram, Oliver Josephs, Ralf Veit, Frank Scharnowski, Rainer Goebel, Niels Birbaumer, Ralf Deichmann, and Klaus Mathiak. 2007. "Real-Time Functional Magnetic Resonance Imaging: Methods and Applications." *Magnetic Resonance Imaging* 25 (6): 989–1003. doi:10.1016/j.mri.2007.02.007.

West, Brady T., Kathleen B. Welch, and Andrzej T Galecki. 2007. *Linear Mixed Models: A Practical Guide Using Statistical Software*. *2003*. Chapman & Hall/CRC Taylor & Francis Group.

Weyand, Sabine, and Tom Chau. 2015. "Challenges of Implementing a Personalized Mental Task near-Infrared Spectroscopy Brain-Computer Interface for a Non-Verbal Young Adult with Motor Impairments." *Developmental Neurorehabilitation* In Press.

Weyand, Sabine, Larissa Schudlo, Kaori Takehara-Nishiuchi, and Tom Chau. 2015. "Usability and Performance-Informed Selection of Personalized Mental Tasks for an Online near-Infrared Spectroscopy Brain-Computer Interface." *Neurophotonics* 2 (2): 025001. doi:10.1117/1.NPh.2.2.025001.

Weyand, Sabine, Kaori Takehara-Nishiuchi, and Tom Chau. 2015a. "Correlates of near-Infrared Spectroscopy Brain-Computer Interface Accuracy in a Multi-Class Personalization Framework." *Frontiers in Human Neuroscience* 9 (536). doi:10.3389/fnhum.2015.00536.

Weyand, Sabine, Kaori Takehara-Nishiuchi, and Tom Chau. 2015b. "Weaning off Mental Tasks to Achieve Voluntary Self-Regulatory Control of a Near-Infrared Spectroscopy Brain-Computer Interface." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 23 (4): 548–61.

Weyand, Sabine, Kaori Takehara-Nishiuchi, and Tom Chau. 2015c. "Exploring Methodological Frameworks For A Mental Task-Based Near-Infrared Spectroscopy Brain-Computer Interface." *Journal of Neuroscience Methods*, no. 254: 36–45.

Wickelgren, Wayne A. 1977. "Speed-Accuracy Tradeoff and Information Processing Dynamics." *Acta Psychologica* 41 (1): 67–85. doi:10.1016/0001-6918(77)90012-9.

Wolf, Martin, Ursula Wolf, Vlad Toronov, Antonios Michalos, L.Adelina Paunescu, Jee Hyun Choi, and Enrico Gratton. 2002. "Different Time Evolution of Oxyhemoglobin and Deoxyhemoglobin Concentration Changes in the Visual and Motor Cortices during Functional Stimulation: A Near-Infrared Spectroscopy Study." *NeuroImage* 16 (3): 704–12. doi:10.1006/nimg.2002.1128.

Wolpaw, J R, D J McFarland, G W Neat, and C a Forneris. 1991. "An EEG-Based Brain-Computer Interface for Cursor Control." *Electroencephalography and Clinical Neurophysiology* 78 (3): 252–59. http://www.ncbi.nlm.nih.gov/pubmed/1707798.

Wolpaw, J R, D J McFarland, and T M Vaughan. 2000. "Brain-Computer Interface Research at the Wadsworth Center." *IEEE Transactions on Rehabilitation Engineering* 8 (2): 222–26. http://www.ncbi.nlm.nih.gov/pubmed/12899275.

Wolpaw, J. R., D. Flotzinger, G. Pfurtscheller, and D. J McFarland. 1997. "Timing of EEG-Based Cursor Control." *Journal of Clinical Neurophysiology* 14 (6): 529–38. http://www.ncbi.nlm.nih.gov/pubmed/9458060.

Wolpaw, Jonathan R, and Dennis J McFarland. 2004. “Control of a Two-Dimensional Movement Signal by a Noninvasive Brain-Computer Interface in Humans.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (51): 17849–54. doi:10.1073/pnas.0403504101.

Wolpaw, JR, N Birbaumer, W.J. Heetderks, D.J. McFarland, P.H. Peckham, G. Schalk, E. Donchin, L.A. Quatrano, C.J Robinson, and T.M. Vaughan. 2000. “Brain-Computer Interface Technology: A Review of the First International Meeting.” *IEEE Transactions on Rehabilitation Engineering* 8 (2): 164–73. http://www.ocf.berkeley.edu/~anandk/neuro/BCI Overview.pdf.

Xiong, J, S Rao, P Jerabek, F Zamarripa, M Woldorff, J Lancaster, and P T Fox. 2000. “Intersubject Variability in Cortical Activations during a Complex Language Task.” *NeuroImage* 12 (3): 326–39. doi:10.1006/nimg.2000.0621.

Yang, Hongyu, Zhenyu Zhou, Yun Liu, Zongcai Ruan, Hui Gong, Qingming Luo, and Zuhong Lu. 2007. “Gender Difference in Hemodynamic Responses of Prefrontal Area to Emotional Stress by near-Infrared Spectroscopy.” *Behavioural Brain Research* 178 (1): 172–76. doi:10.1016/j.bbr.2006.11.039.

Yu, Lei, and Huan Liu. 2003. “Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution.” In *Proceedings of the Twentieth International Conference on Machine Learning*, 2:856–63.

Zagar, Robert John, Joseph W. Kovach, Kenneth G. Busch, Michael D. Zablocki, William Osnowitz, Jonas Neuhengen, Yutong Liu, and Agata Karolina Zagar. 2013. “Ammons Quick Test Validity Among Randomly Selected Referrals.” *Psychological Reports* 113 (3): 823–54. doi:10.2466/03.04.PR0.113x29z0.

Zephaniah, Phillips V., and Jae Gwan Kim. 2014. “Recent Functional near Infrared Spectroscopy Based Brain Computer Interface Systems: Developments, Applications and Challenges.” *Biomedical Engineering Letters* 4 (3): 223–30. doi:10.1007/s13534-014-0156-9.

Zhu, H., M. Liu, H. Shu, H. Zhang, and L. Luo. 2010. “General Form for Obtaining Discrete Orthogonal Moments.” *IET Image Processing* 4 (March): 335. doi:10.1049/iet-ipr.2009.0195.