Designing Interactive Applications using Active and Passive EEG-based BCI Systems

by

Chi Thanh Vi

A thesis submitted to the University of Bristol
in accordance with the requirements of the degree of Doctor of Philosophy
in the Faculty of Engineering, Department of Computer Science

May 2014

Word count: 53253
Abstract

A brain computer interface (BCI) is a communication system that allows users to control computers or external devices by detecting and interpreting brain activities. The initial goals of BCI were to help severely disabled people, such as people with "locked-in" syndrome, to communicate with the outside world by interpreting their brain signals into corresponding external commands. Nowadays, state-of-the-art BCIs, especially using Electroencephalography (EEG), bring benefits to normal and healthy computer users in a way that enriches their experiences of everyday Human Computer Interaction (HCI). Although EEG may be used in the same manner of continuous control and communication, it has been extended to assist and measure the inner states of users in a more passive way. Because of this, a new categorization of BCI systems has been proposed, dividing BCI applications in general and EEG-based systems in specific into active, reactive, and passive BCI.

This thesis focuses on how portable and commodity EEG headsets can benefit the majority of HCI users with their limited capabilities, in comparison to clinical and expensive headsets. Our investigations focus on active and passive EEG-based BCI systems. We first investigate about how to use task engagement as an additional input besides traditional input methods in the context of active BCI. We then move forward to passive use of BCI by using task engagement to evaluate an application while the user is taking part in an interaction. We further extend our investigation to Event-Related Potentials where in particular Error-Related Negativity is used to detect users’ error awareness moments. We show that using EEG signals captured by Emotiv headsets, moments of users’ error awareness (or Error Related Negativity - ERN) can be detected on a single trial basis. We then show that the classification rates are sufficient to benefit HCI in single user. Next, we show ERN patterns can be detected in observation tasks where it not only appears in the observers' EEG, but also shows an anticipation effect in collaborative settings. Based on the results, we propose different scenarios where task designers can employ these results to enhance interactive applications, combining with popular HCI settings and input methods.
Declaration of Authorship

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Signed:

Date:
This thesis was advised by:

Professor Sriram Subramanian
Department of Computer Science
University of Bristol
Bristol, UK

It was examined on 10th July 2014 by:

Professor Robert Jacob
Dept. of Computer Science
Tufts University
Massachusetts. U.S.A
(as external examiner)

&

Professor Mike Fraser
Department of Computer Science
University of Bristol
Bristol, UK.
(as internal examiner)
I have never met a man so ignorant that I couldn’t learn something from him.

- Galileo Galilei
Acknowledgements

It is said that a PhD is a special journey, not just a degree. To me, this journey was challenging but joyful and eventful. I started my PhD with no prior publications, not much experiences of conducting research. However along the journey, I have achieved many accomplishments, including having my little family. I remember what my advisor said at the very first day of my PhD: “A PhD is a long journey. You may have to stick with something for 3 years. And you may want to give up several times when things don’t work out as you want”. Now looking back at this journey, I completely agree. It was much easier to just run away when you stuck. Fortunately, I have lots and lots of support from many people around. I truly appreciate for their help during my PhD course.

I am lucky to have continuous and unconditional support from my family and loved ones. My father, Vi Văn Liên, and my mother, Trần Thị Lan, always be there and make sure that I have more than enough resources to satisfy my needs to pursue my PhD degree. To my wife, Tú Anh, who has been constantly supporting and being my reliable ground, love you always. To my little angel, “Mango” Xoài, you are the gift of my life. I hope my achievements so far will make all of you proud. To my beloved grandmas, grandpas, uncles, aunties, cousins and in-laws, thank you for your prayers and well-wishes. To my little sisters, Mai Hường, Giang, and Season, thank you for always are supportive and cheerful.

It is said that a journey is made most memorable by quality companions. True to this, I was lucky to have a motley crew of colleagues to thank for making the four years a very memorable experience. I wish to thank “Mr. Abe” Abhijit Karnik, “Izzy” Izdihar Jamil, Benjamin “Benny Ben” Long, Jason Alexander, Mark Marshall, Anne Roudaut, along with other current and past members of the BIG lab. I also want to express my gratitude to Kitamura sensei, Kazuki Takashima, Hitomi Yokoyama, “JJ” Jeongjoo Hong and his wife Fumi Chiba, along with other current and alumni members of the ICD lab. You made every moments of my time in Sendai, Japan great, eventful, and with wonderful experiences. I also thank Dr. Rafal Bogacz and Professor Mike Fraser, my internal reviewers. Your helps during my PhD and your comments for my reviews allow me to make this far. I also thank Dr. Jade Ngoc Thai and other members at CRICBristol, with their help to make sure my experiment runs smoothly. I thank Dr. Nick Yeung at University of Oxford for providing early guidance to my research topic. I’d still like to thank Heka a.k.a my trusty computer which make all the work possible.

Last but not the least, I cannot say thank enough to my advisor, Professor Sriram Subramanian. He is a true visionary who always provide me guidance, creativity, and inspiration. I lost count how many times he diverted me back to the HCI field with his “hand-holding” support. His responsive inputs and encouragement allowed to drive me this far. Not only my research has grown under his supervision, I have too. He has pushed me to excel and I hope this thesis answers his trademark question, “Is there any result yet?” in a more satisfactory way.
# Table of Contents

Chapter 1. Introduction ........................................................................................................... 1

1.1. Problem Space ............................................................................................................... 1

1.1.1. ERP and ERP ........................................................................................................... 2

1.2. Vision ............................................................................................................................ 3

1.3. Research goals .............................................................................................................. 4

1.3.1. Methodology ........................................................................................................... 5

1.4. Thesis Organization ..................................................................................................... 6

Chapter 2: Background .......................................................................................................... 11

2.1. The human brain .......................................................................................................... 11

2.1.1. Frontal lobe ............................................................................................................ 12

2.1.2. Parietal lobes ......................................................................................................... 12

2.1.3. Occipital lobe ....................................................................................................... 12

2.1.4. Temporal lobes ...................................................................................................... 13

2.2. The Anterior Cingulate Cortex .................................................................................. 13

2.3. Measuring brain activity ............................................................................................. 14

2.3.1. Different measuring techniques: ............................................................................ 14

2.3.2. Electroencephalography (EEG) ............................................................................. 17

2.3.3. Active and passive Brain Computer Interface ....................................................... 21

2.4. Interpreting brain activities in emotions and cognitive states .................................... 21

2.4.1. Meditation ............................................................................................................. 22

2.4.2. Task engagement .................................................................................................. 22

2.4.3. Evaluate interactive systems using different types of emotions ........................... 23

2.5. Event-related potentials (ERP) ................................................................................... 23

1.5. Connections between chapters ..................................................................................... 9

1.6. Contributions ................................................................................................................. 9

1.6.1. Research contributions ........................................................................................... 9

1.6.2. Publications ............................................................................................................ 10

Chapter 3: Quantifying Task Engagement for use in Gaming Applications .................... 7

Chapter 4: Measuring task engagement to evaluate interactive tasks ............................. 7

Chapter 5: Error-Related Negativity for Single User Interactive Applications ............. 7

Chapter 6: Error Related Negativity in Collaborative Applications ............................... 8

Chapter 7: Discussion and conclusions ............................................................................. 8

x
2.5.1. P300 evoked potentials ................................................................. 24
2.5.2. Error-related negativity (ERN) ....................................................... 24

2.6. Data acquisition and processing ....................................................... 27
  2.6.1. Hardware .................................................................................. 27
  2.6.2. Evaluate classifiers using the Receiving Operating Characteristic .... 29

2.7. Summary ......................................................................................... 29

Chapter 3. Quantifying Task Engagement for use in Gaming Applications ...... 31
  3.1. Potentials of commodity EEG headsets .......................................... 31
  3.2. Related Work ............................................................................... 33
  3.3. Factors Influencing the Design of a System using Engagement for Control ................................................................................. 34
    3.3.1. Engagement Discretization .......................................................... 35
    3.3.2. Mapping Method ...................................................................... 35
    3.3.3. Selection Technique .................................................................. 36
    3.3.4. Visual Feedback ........................................................................ 36
    3.3.5. Calibrating engagement .............................................................. 36

  3.4. Experiment Setup and Signal Processing ............................................. 37
    3.4.1. Experiment Setup ...................................................................... 37
    3.4.2. Signal Processing ...................................................................... 38

  3.5. Experiment 1: Identifying levels of control ........................................ 39
    3.5.1. Experimental Procedure and Interface ....................................... 40
    3.5.2. Study 1A .................................................................................. 41
    3.5.3. Study 1B .................................................................................. 43
    3.5.4. Discussion: Study 1A and 1B ..................................................... 44

  3.6. Experiment 2: Engagement-based distance control ............................. 44
    3.6.1. Experimental Design and Procedure ......................................... 45
    3.6.2. Participants .............................................................................. 45
    3.6.3. Results and Discussion .............................................................. 45

  3.7. Experiment 3: Distractions & multi-tasking in a game setting ............... 47
    3.7.1. Design ..................................................................................... 47
    3.7.2. Experimental Interface and Task ............................................... 47
    3.7.3. Experimental Design and Procedure ......................................... 48
    3.7.4. Participants .............................................................................. 48
    3.7.5. Results & Discussion .............................................................. 48

  3.8. Discussion, Implications, and Applications in Game Design ................ 49
    3.8.1. Continuous vs. Discrete Input Control ...................................... 49
3.8.2. Error Resilience in BCI Game Design ................................................... 50
3.8.3. Lessons for Designers ................................................................. 50
3.9. Summary ..................................................................................... 51

Chapter 4. Measuring task engagement to evaluate interactive tasks .......... 52
4.1. The design of D-FLIP ...................................................................... 53
4.2. Algorithm Overview of Dynamic Display of Photos ............................ 54
  4.2.1. Principles of Photograph Arrangement ........................................ 54
  4.2.2. Behaviours of Photographs and Performance .............................. 57
4.3. Discussions ................................................................................. 59
4.4. Performance Evaluation .................................................................. 60
4.5. Measure and process EEG signals to calculate task engagement .............. 61
  4.5.1. The Effect of Animation on Users’ Interest .................................... 61
  4.5.2. Evaluating using Neural Signals ................................................ 62
  4.5.3. User evaluation of d-flip ........................................................... 63
4.6. Discussion .................................................................................... 68
4.7. Summary ..................................................................................... 69

Chapter 5. Error-Related Negativity for Single User Interactive Applications ...... 71
5.1. Introduction .................................................................................. 71
5.2. Error Related Negativity (ERN) .................................................... 73
5.3. Online ERN Detection .................................................................... 74
5.4. Experiment 1: Flanker Task ............................................................ 76
  5.4.1. Signal Processing ....................................................................... 76
  5.4.2. Results ..................................................................................... 77
  5.4.3. Discussion ................................................................................ 79
5.5. Experiment 2: Buttons selection ...................................................... 79
  5.5.1. Task: Button selection ............................................................ 80
  5.5.2. Design and procedure ................................................................ 81
  5.5.3. Data Collection and Analysis .................................................... 81
  5.5.4. Result ..................................................................................... 82
  5.5.5. Discussion ................................................................................ 83
5.6. Experiment 3 ................................................................................ 83
  5.6.1. Task and Technique .................................................................. 84
  5.6.2. Design and Procedure ............................................................ 85
  5.6.3. Result ..................................................................................... 85
List of Figures

Figure 1. Left view of the brain ................................................................. 11
Figure 2. The 3Tesla Siemens Magnetom Skyra MRI scanner at Clinical Imaging Research Centre Bristol................................................................. 15
Figure 3. fNIR 1100 Imager with 4 photo-emitters and 10 photo-detectors (Biopac system) ................................................................. 15
Figure 4. A user is performing a task in the MEG machine ................................................................. 16
Figure 5. The first human recorded electroencephalography [139] ................................................................. 17
Figure 6. Locations of the 10-20 system, as standardized by the American Electroencephalographic Society [121] ................................................................. 18
Figure 7. Different type of EEG frequency bands examples [121] ................................................................. 19
Figure 8. Location of the anterior cingulate cortex (ACC – in red) in the human brain .............................................................................. 25
Figure 9. An example of ROC graph .............................................................................. 29
Figure 10. A user wearing the Emotiv EEG headset in the final study setup that combines engagement control and traditional input ................................................................. 32
Figure 11. Three objects of the 3CVT. Objects appeared without captions in the experiment. .............................................................................. 38
Figure 12. Signal processing procedure .............................................................................. 38
Figure 13. Bull’s eye (left) and Linear (right) visualizations. The yellow circle (square) shows the user’s state, the light green circle (square) shows the start state, the bright green circle/square at the end state, and the red line/circle shows the cursor .............................................................................. 40
Figure 14. Task completion time in sec (left) and average number of false targets selected (right) for each state (horizontal axis) and visualizations (B / L = Bull’s-eye / Linear Visualization) for Study 1A .............................................................................. 42
Figure 15. Task completion time in sec (left) and false targets selected (right) over distances for Study 1A .............................................................................. 42
Figure 16. Task completion time in sec (left) and false targets selected (right) over number of states (horizontal axis) and visualizations (B / L = Bull’s-eye / Linear Visualization) for Study 1B .............................................................................. 43
Figure 17. Average max(EI) and min(EI) measured in the first two tasks over all participants of study 1A and 1B .............................................................................. 44
Figure 18. Mean task completion time for experiment 2; trend line shows linear correlation between time (vertical axis, in seconds) & distance (horizontal axis) .............................................................................. 46
Figure 19. Mean false targets selected (vertical axis) for each distance (horizontal axis) with standard error bars for experiment two ........................................................................................................... 46
Figure 20. Visualization of the avatar (car), fireball (red circle), and engagement levels in case of not-matched (left) and matched (right) ........................................................................................................... 47
Figure 21. Number of false targets and mean stop duration for Experienced (E) and Novice (N) users ..................................................................................................................................................... 49
Figure 22. Examples of photo arrangements by Geotags (a); find someone’s photos (b); and photo corrected by meta-data (c) ........................................................................................................................................ 52
Figure 23. Conceptual behaviours of photographs ........................................................................................................................................................................ 56
Figure 24. Geometric packing of photos without the rotation ........................................................................................................................................ 58
Figure 25. Geometric packing of photos with the rotation ........................................................................................................................................ 58
Figure 26. Sequence when size of the displaying window is changed ........................................................................................................................................ 58
Figure 27. Examples of geometric packing where a photo is focused by overlaying a pointer (a); and semantic mapping where photos are arranged by colour and interest similar colour photos are collected (b) ........................................................................................................................................ 59
Figure 28. Transition of frame rate ..................................................................................................................................................................................... 60
Figure 29. Storytelling mode with D-FLIP (left) and Windows Explorer (right) ........................................................................................................................................ 64
Figure 30. An example of a narrator’s task engagement changes in time in one session (left); and Task engagement of each narrator (right) ........................................................................................................................................ 67
Figure 31. Task engagements of listeners in the Story Preparation task (left) and Story Telling task (right) between Windows Explorer and D-FLIP ........................................................................................................................................ 67
Figure 32. A user performing a Superflick [154] while wearing an off-the shelf EEG headset to see if ERN detection can aide in an object selection task ........................................................................................................................................ 72
Figure 33. Online Single Trial ERN Detection Procedure ........................................................................................................................................ 75
Figure 34. Coefficient matrix extractor ..................................................................................................................................................................................... 75
Figure 35. Examples of single trial ERN for two cases: incorrect (a) and correct (b). The x axis is number of samples (1 sec = 128 samples). The y axis is amplitude (uV). The red line is the key press moment (sample number 127) ........................................................................................................................................ 77
Figure 36. Classifier accuracy on a single trial basis ..................................................................................................................................................... 78
Figure 37. EEG signals at channel AF4 with number of samples in the x-axis (1 sec = 128 samples) and uV in the y-axis. (a) Correct Epochs and (b) Incorrect Epochs. The vertical red line is the moment when key was pressed ........................................................................................................................................ 78
Figure 38. ROC curve for 4 channels: AF4, F3, F8, and FC6 ..................................................................................................................................................... 78
Figure 39. AUC for different channels ..................................................................................................................................................................................... 79
Figure 40. User interface of the task at the moment of (a) start trial and (b) cursor is near the desired button. ............................................................... 80
Figure 41. EEG signals around moments where a participant made an incorrect decision (a, c) and correct decision (b, d) at channel F3 (top) and FC5 (bottom). The red lines mark the moment that the mouse button was clicked. ........................................................................ 82
Figure 42. Single trial classification accuracies ........................................................................ 83
Figure 43. Accuracies of Superflick with different ERN detection accuracies .................. 86
Figure 44. Observation of Flanker Task with the executor (left) and the observer (right) .......... 93
Figure 45. Mean classifier accuracy (left) and AUC (right) for different channels. .................. 94
Figure 46. Average EEG signals at channel F3 (top) and F7 (bottom).................. 94
Figure 47. ROC curves for channels F3 and F7................................................. 95
Figure 48. Tabletop Flanker task a) close, b) far layout. .................................................. 98
Figure 49. Averaged signals of the ‘close’ layout at Cz.................................................. 99
Figure 50. Topographic distribution of correct trials (top) and incorrect trials (bottom) in the ‘close’ layout at the intervals: key press (left), 100ms after (middle), and 200ms after (right).100
Figure 51. Averaged signals of the ‘far’ layout at Cz.................................................. 100
Figure 52. Topographic distribution of correct trials (top) and incorrect trials (bottom) in the ‘far’ layout at the intervals: 50ms before key press (left), key press (middle), and 50ms after key press (right). .................................................................................... 101
Figure 53. Classification rates of close layout (left) and far layout (right)......................... 101
Figure 54. AUC for close (left) and far layouts (right).................................................... 101
Figure 55. Averaged signals of the close layout at AF3. The vertical line is the button touched moments.................................................................................................................. 103
Figure 56. Average signals of the far layout at F3. The left vertical line is the averaged hand lift-off moments, the right vertical line is the button touched moments............................................. 104
Figure 57. Classification rates of experiment 3 ............................................................. 104
Figure 58. BCI tabletop study example ................................................................. 133
List of Tables

Table 1. 10-20 international system acronyms explanation ............................................. 18
Table 2. Average task engagement for narrators and listeners............................................. 66
Table 3. Mean % of unsuccessful trials with the two indication light.................................... 86
List of Publications


Chapter 1. Introduction

1.1. Problem Space

A Brain Computer Interface (BCI) is a communication system, either hardware or software, which allows users to control or interact with an external devices using their brain signals, without using any peripheral muscles. A popular brain sensing technique in BCI is Electroencephalography (EEG). It is the recording of the electrical fields produced by neuronal activity [55]. EEG has been used to assist users to communicate with and control external machines such as computers and assistive appliances.

Initially, BCI was mainly designed to help people with several motor disabilities (such as patients with locked-in syndrome [112] where the patients are aware and awake but cannot move or communicate verbally). In this example, the BCI system can be used to provide a communication medium and consequently improve their quality of life. However, such applications have remained unattractive for several reasons. First, existing research, which employs brain activity, has usually been restricted to clinical settings and focused on neurological disorders investigation [137]. Second, design of the BCI systems was complex due to the high variability in the captured brain signals, which leads to limited input bandwidth of the information. Third, the BCI systems benefit users the most when their brain signals are processed right after being captured, in real-time. This technology requires extremely expensive equipment or more advanced signal processing techniques [192], which limits the use of BCI in everyday life, for wide population of users.

Recently with the advancement of technology in this field, affordable EEG kits such as the Emotiv EPOC and OCZ Neural Impulse Actuator, NIA have become available, allowing designers to explore BCI interaction techniques further. For example, applications using the EEG signals captured through non-invasive electrodes placed on a user’s scalp range from gaming [88] to robot control [8]. Researchers have mainly demonstrated the capability of these controllers through specific point-designs in applications such as user task classification [113] to controlling a wheelchair [114].

In these applications, EEG signals are commonly used to capture the affective state (such as relaxation [88]) of the user which can then be mapped to controlling virtual [114] or physical objects [88]. A range of possibilities with sensed parameters have been further explored,
ranging from arousal [52] to engagement [27] and relaxation [88]. However, there exist only a limited number of generic design principles that can guide application developers in mapping a sensed parameter, such as the level of engagement, to virtual control. Answers to some of the fundamental control issues such as “how many discrete levels of engagement can a user comfortably control?” or “is it easier for users to control a parameter by becoming more engaged than less engaged?” are largely unexplored. Answers to such questions can allow designers to break away from the confines of point-designs to more generic forms of mappings of EEG signals to virtual control.

1.1.1. ERP and ERP

Another popular usage of EEG in BCI is to detect and design interactive applications using Event-Related Potentials (ERPs). ERPs are captured brain responses resulting from specific events such as sensory, cognitive, or motor [119]. To date, most of these applications have been based on using P300 (a peak elicited at about 300ms after an event onset) which can be considered as a combined index of both attention and memory [113]. Another ERP that could be used but is often ignored in interactive applications is the Error-related negativity (ERN).

An ERN is a form of an Event Related Potential (ERP) that can be triggered in the brain when users either make a mistake or the application behaves differently from their expectations. This pattern is produced in a person’s brain when they are aware of the obvious error(s) that they have made; either through system feedback or individual realization [68]. For example, ERN would be produced when pressing the LEFT key while intending to press the RIGHT key in a multiple choice reaction time task. It also appears, but with lower amplitude when users are confused about the decision that they have made [47]. Usually these ERN signals appear and peak within 150ms of the committed action [73].

However it is difficult to detect a clear ERN pattern due to noisy EEG signals and lack of effective real-time algorithms. Most research in detecting ERN is focused on being able to detect it over an average of multiple trials relying on offline methods (see [47, 130] for examples). As ERN is known to appear when using multiple-choice reaction time tasks, most research is done on Flanker task [61]. Flanker task is a visual experiment where the participant has to respond to a central and directed symbol that is surrounded by distracting symbols. There are very few ERN detection algorithms that work in real-time as well as on multiple-choice RT tasks that are not Flanker tasks.
In addition, most examples of ERN to date have focused on detecting the signal by recording signals from the executer of the task. Recent endeavours [50, 132, 182] suggest that ERN signals even appear while observing another user making errors. These studies confirm the existence of negative potential within 250ms after the event onset and analysis of the signals' origin confirm that they are ERNs. This is useful in scenarios where the executer is not aware of their error but that error is spotted by an observer or supervisor. In current studies, the observer did not have much opportunity to anticipate the executer’s actions and often relied on the answer displayed. In HCI scenarios, the executer’s actions may become visible to the observer even before the action is committed. For example, in collaborative tabletops, an observer has awareness of an executer’s actions through their reaching gesture [147].

Moreover, most ERN experiments are performed using expensive, non-portable sensing devices such as NeuroScan systems from NeuroScan™, g.BCIsys from g.tec™ or ActiveTwo from Biosemi™. These devices have the benefits of not only more sensing channels (up to 256 electrodes therefore can cover all channels in the 10-20 international system), but also higher and user selectable sampling frequencies of up to 16 kHz. These systems also introduce less noise in the signals because of better sensing electrodes and integrated amplifier/ converter modules.

In comparison, the off-the-shelf EEG headsets have at best 14 channels, a sampling frequency of about 1 kHz and introduce more noise due to wireless transfer of EEG signals. Due to these limitations it is not clear if these devices can capture ERN patterns. However, due to their price and portability factors their use is being increasingly explored within HCI. As a consequence, it is interesting to investigate the appropriateness of these off-the-shelf EEG headsets in detecting ERNs and also the effectiveness of real-time algorithms in doing so on a single trial and in real-time.

1.2. Vision

We envision a future where a user interacts with a device, this device can give appropriate information or guidelines based on the detected brain signals. The brain signals captured from the interacting user can be directly controlled by them as an additional input with a reasonable input bandwidth compared with traditional input devices such as mouse, keyboard, joysticks, etc. In addition to this, the captured neuronal signals can also help to understand users’ underlying thoughts. This can be done passively so that the machine can assist the users based on the context of the on-going task without interrupting the workflow. For example, the amount of
work assigned to a user can be adaptively adjusted according to that user’s workload. Another example is the system providing appropriate information or guidelines based on users’ cognitive states such as confusion, error awareness, and different emotions (happiness, fear, sadness, etc.). In addition, the devices that capture the brain signals are portable, non-obstructive thus making users feel comfortable when using it.

The scenarios described certainly raise interesting questions and challenges. The main challenge is that the BCI-based systems are unattractive to users due to low input bandwidth, speed, and task restrictions as well as restricted interactive gestures. Moreover, the rates for the accurate detection of users’ underlying cognitive states are not attractive enough compared to other common input devices, which are typically operated by hand gestures. Although it is shown that users’ cognitive states can be detected, available guidelines on how to harness them in order to enrich interactive tasks remain unsolved.

1.3. Research goals

We aim to bring BCI closer to the mass users. Our first goal is to produce generic design principles for controlling EEG sensed parameters. Task engagement is chosen as the parameter (hereafter “engagement”) to achieve this first goal. Task engagement is the effortful concentration used to strive towards task goals [124]. Commercial EEG kits (such as the Emotiv EPOC) can sense levels of engagement in real time giving interaction designers the ability to exploit this fundamental mental state as a control mechanism for various interactive activities. Previous research has derived methods to successfully measure engagement using EEG [42, 148], but the low-level questions for interactive control listed above remain unanswered. In addition, we aim to employ engagement as a tool to evaluate interactive tasks. Therefore, we will first develop a visualized interactive task as an example of common interactive applications. Following this we used engagement to evaluate the application while it is being used compared to another application.

For our second goal, we aim to investigate the capabilities of off-the-shelf headset in detecting Error-Related Negativity for assisting interaction design. This is because if ERN signals can be detected during an interactive task they can be used in detecting and correcting errors or in augmenting users’ experiences in those activities. An ERN detection module provides another medium for HCI designers to access users' intentions, which is intuitive and originates directly from the users’ brain. This has great potential in many types of interactive application such as in
gaming, spatial navigation tasks and aiding object selection. For example, when ERN is detected the system can prompt the user to check if the selection is the intended target or not.

For our third and final goal, we aim to investigate the effect of anticipation in an observing task where the outcome of an action is revealed before that action is committed. This has not been investigated with ERN in an observing task before. This can have many applications in pair-programming, collaborative tabletops and emergency response applications. We first investigate these anticipation effects with an expensive clinical EEG system to demonstrate that the signal is present and can be detected. Following this we investigate if an Emotiv headset has the similar detection capability.

1.3.1. Methodology

Our approach to answer the research questions is different from the neuroscience approach. Our investigations are based on the existence of users’ cognitive states confirmed by earlier neuroscience related works. Therefore our work focuses on how to detect and harness these states in the appropriate HCI scenarios. Our approach includes the design space exploration, task designing, and user evaluations.

a. Design space exploration

Human cognitive states, ERP and their triggering conditions have been widely investigated in the neuroscience fields. For the designing of BCI-based interactive tasks, we explore the possibility of the different discovered brain patterns and states to identify which one is the simplest to implement and at the same time leads to further investigations. This involved identifying and categorizing what have been discovered in the neuroscience field, then harness them to benefit HCI. This is done using our existing hardware tools and restrictions about applying them for majority of normal and healthy users.

b. Task designing

Once the features were chosen to investigate their usefulness for HCI, the task-designing phase required carefully designed interactive tasks. Indeed, while the conditions that trigger the investigated features should be preserved, these tasks should be pushed towards HCI interactive tasks (e.g. with different input devices, with more or longer gestures during task performance and with the associated generated noises). There are several steps taken between capturing the brain signals and using them as an input for the system, which are described later in this thesis.
c. User evaluations

We measured the EEG signals using the Emotiv headsets to calculate task engagement and to detect ERN pattern, which had not been done before. We then evaluated how well the used algorithm performed compared to other existing studies conducted with other devices (usually more expensive) and tasks (usually more restricted).

For chapters 3 and 4, we aimed to detect task engagement for use as either an input method or an evaluation tool. We performed series of user evaluation to investigate this step by step through task driven experiments. The results of these studies confirmed both the usefulness of using task engagement to enrich and evaluate HCI and the capabilities of off-the-shelf headsets (e.g. Emotiv EEG headset).

For chapters 5 and 6, we first replicate user studies that were used in neuroscience studies to detect ERN pattern. In this first step, we aimed to detect ERN patterns using Emotiv, a commodity headset, on a single trial basis. Once the results confirmed the possibilities, we proceeded to further user evaluations to investigate different scenarios where ERN is beneficial.

1.4. Thesis Organization

The structure of this thesis is organized following the works through my PhD. At first, the general related works are described. Following are several chapters covering each work.

Chapter 2: Related Works

This chapter gives a comprehensive overview of the related work pertaining to the areas of brain sensing techniques, measuring tasks engagement, and ERN. We begin with presenting the overview structure of the human brain and how its neural activities are being measured using different sensing techniques, especially EEG. We then discuss how EEG has been used to interpret human’s affective states such as meditation and task engagement.

In addition, different types of ERPs are also be covered with an extensive look into ERN such as source localization of ERN, ERN in a single user scenario, ERN in observation tasks as well as different attempts in detecting ERN for interaction designs. Background overview on EEG data acquisition and processing are described with more details on signal processing, feature selection, statistical classifier and how to evaluate them. Finally, we discuss regarding the active, reactive, and passive BCI systems. Our investigations in this thesis cover both the active and the passive BCI-based system. The active BCI system is where the brain signals are used for
controlling directly while the passive BCI-based system is where the brain signals are triggered without the purpose of voluntary control.

**Chapter 3: Quantifying Task Engagement for use in Gaming Applications**

The goal of our investigation is to design an active BCI-based system where neuronal signals that are captured from users can then be used directly as an input in applications (e.g. in gaming applications). Therefore in this chapter, we aim to answer to some of the fundamental control issues such as “how many discrete levels of engagement can a user comfortably control?” or “is it easier for users to control a parameter by becoming more engaged than less engaged?” which are largely unexplored. We implemented three experiments where the first two were to generate principles for one-dimensional selection in the absence of environmental or task distractions. The third was an ecologically valid experiment for controlling gaming parameters through levels of engagement.

**Chapter 4: Measuring task engagement to evaluate interactive tasks**

In this chapter, we move forward to investigate how to harness task engagement for a passive BCI-based system. Consequently, we present D-FLIP as an example of interactive application. It is a novel algorithm that dynamically displays a set of digital photos using different principles for organizing them. A variety of requirements for photo arrangements can be flexibly replaced or added through the interaction and the results are continuously and dynamically displayed. As a consequence, the global layout of all photos is automatically varied. We first present examples of photograph behaviours that demonstrate the algorithm and then investigate users’ task engagement using EEG in the context of story preparation and telling. This is organized to show how EEG can be used to evaluate an interactive application, as D-FLIP as an example.

**Chapter 5: Error-Related Negativity for Single User Interactive Applications**

We continue to explore passive BCI systems where EEG signals are captured in background to potentially enrich the interaction. We investigated the usage of commodity EEG headsets through ERP, a brain pattern elicited in certain task conditions. Following this, we describe in this chapter an online single trial ERN detection technique that is verified using data acquired from the frontal-central cortex of the human brain. We then show that we can detect ERN online on a single trial in an object selection task. This demonstrates the abilities of harnessing ERN in interactive applications in office conditions. Both our experiments show that we can detect ERNs using the Emotiv Headset with an accuracy of up to 70%.
These rates are indicative of the type of accuracy one can expect from off-the-shelf EEG sets. Improvements in learning techniques can improve this accuracy but it is unlikely that detection accuracy will reach 100%. It is therefore expected that with high detection accuracy users might become over-resilient on a system detecting ERNs which can increase the cost of recovery from the error. In order to examine this issue we conducted a final study where we compared users’ error-rates when performing Superflick [154], a modified pointing task, with different potential ERN success rates. The results of the experiment show that ERN detection rates of 65 to 80% are acceptable for interactive applications. We finally discuss the implications of our results for interactive applications.

Chapter 6: Error Related Negativity in Collaborative Applications

Following the investigations of ERN for single user settings in chapter 5, we investigate the effect of anticipation in an observing task where the outcome of an action is revealed before that action is committed. This has not been investigated with ERN in an observing task before. Our findings can then be used in many applications in pair-programming, collaborative tabletops and emergency response applications. We investigate these anticipation effects with an expensive clinical EEG system to demonstrate that the signal is present and can be detected. Following this we investigate if an Emotiv headset has the similar detection capability.

We start by repeating an experiment from van Schie et al. [182] with an Emotiv headset to demonstrate that Observer ERN can be detected in a single trial and online basis, with an accuracy up to 64%, using a commodity headset. After that, we investigated the anticipation effect in observing tasks when one person observes another person committing errors. Our results show that there are ERN-like patterns detected in the observer’s EEG about 368ms after the initial movements happened and 55ms before the errors are committed. Following this result, we then show that the Emotiv headset can capture these patterns in the same experiment settings. Finally, we discuss the implications of our results on interactive applications.

Chapter 7: Discussion and conclusions

This chapter summarizes the contributions demonstrated in the previous chapters. We look back at the investigated work and aims to identify how they can be improved on the performance side. We also discuss the implications of our results, obtained from the studies in this thesis, on interactive applications. In addition, guidelines on how to harness task engagement, ERN in single user- and collaborative settings are discussed here. Lastly, this chapter raises questions for future researches based on the presented work throughout the thesis.
1.5. Connections between chapters

This thesis aims to investigate the guidelines and applicability of employing Brain Computer Interface to a majority of computer users. However, it is impossible to cover all aspects of the investigation in this single PhD thesis. Therefore, the first criterion that we follow is to explore areas that are well-researched in the neuroscience literature so that there is a clearly defined starting point for our exploration. The second criterion is to choose ideas that are novel enough to investigate through a PhD thesis. As a consequence, we examine and provide guidelines for HCI designers in the context of both active and passive EEG-based BCI systems. We first explore generic design principles for controlling task engagement as well as using it to evaluate interactive applications. We then aim to detect Error Related Negativity using off-the-shelf EEG headset for assisting interaction design. The chosen investigating parameters are well-researched in neuroscience in term of how to detect and measure them (supporting our first criterion). Moreover, there are little research or design guidelines of how to harness their benefits to enrich interactive applications as well as user experiences (supporting our second criterion).

1.6. Contributions

The research results are the direct outcomes from the research carried out. Additionally, the results were published in major peer-reviewed conferences as papers. We will describe the main contributions of this thesis in the following sections.

1.6.1. Research contributions

In terms of measuring tasks engagement for evaluation and continuous control, we have made the following contributions,

- We present a systematic exploration of parameters that influence mappings of engagement to virtual cursor control.
- We quantify the number of discrete levels that can be controlled via engagement states from EEG input.
- We present a series of guidelines for designers
- We demonstrate the practical value of our mappings in a gaming application.
- We propose an evaluation method using EEG that can be used to evaluate interactive applications.

Additionally, we also present three contributions in detecting Error-Related Negativity for interaction design, which are:
a) We demonstrate that off-the-shelf EEG devices such as the Emotiv headset can measure ERN from channels in the front-central part of the brain.

b) We present that it is possible to detect these ERN signals online (as opposed to offline) from a single trial in a task that is closer to the types of tasks encountered in HCI.

c) We show that detection accuracies in the region of 65 to 80% are sufficient to use these techniques in real-time interactive applications.

More contributions are presented in the context of observing interactive tasks:

a) We demonstrate that off-the-shelf EEG devices like an Emotiv EEG headset can capture ERN in an observing task from channels in the frontal-central part of the brain.

b) We investigate the anticipation effects in collaborative settings demonstrated by the ERN detected in an observer’s EEG signals before the action is committed.

c) We show that these anticipation effects can still be demonstrated using off-the-shelf EEG devices such as the Emotiv EEG headset.

1.6.2. Publications

The following papers were published and presented in peer-reviewed conferences. The list of these papers is organized in the chronological order of their publication:


P2: “D-FLIP: Dynamic & Flexible Interactive PhotoShow” in ACE 2013 [185]

P3: “D-FLIP: Dynamic & Flexible Interactive PhotoShow” in SIGGRAPH Asia E-Tech 2013 [105]

P4: “Error Related Negativity in Observing Interactive Tasks” in CHI 2014 [186]
Chapter 2. Background

This chapter discusses the background information regarding the human brain, its functions, as well as methods to measure the brain activities. Electroencephalography (EEG) is also covered in more detail, as it is one the most widely used methods to sense the neuronal activities for Brain Computer Interfaces (BCIs). The second part of this chapter describes some common affective states where EEG signals are interpreted. The third part of this chapter covers the background information about Event Related Potentials (ERPs) and Error-Related Negativity (ERN), which is one specific type of ERP. Finally, this chapter explains the related signal acquisitions and processing techniques involved when processing ERN/ERPs.

2.1. The human brain

The human brain controls and coordinates movements, behaviours and homeostasis of internal functions such as heartbeat, blood pressure and body temperature. It is also the site of intelligence and reason including cognition, perception, attention, memory and emotion. Overall, the brain contains an estimated number of 86 billion neurons [17], each linked to as many as 10,000 other neurons [122]. The brain consumes a large amount of energy using about 20% of the oxygen absorbed by the lungs and $\frac{3}{4}$ of the blood sugar produced by the liver [86].

Anatomically, the brain consists of two hemispheres with similar weight and volume [176], the cerebrum, and the brain stem. The two hemispheres are connected through the corpus callosum, which is a large collection of neural fibres underneath the cortex [18]. The brainstem contains
the medulla oblongata, pons, midbrain, and diencephalon. It is responsible for functions such as controlling arousal, alertness, and consciousness [164]. The thalamus, consisting of a large collection of nuclei, transmits nerve impulses between the cerebellum and the basal ganglia that control muscle movements [140].

In addition, on the outer surfaces of the hemispheres is the cerebral cortex. It consists of a large collection of brain neurons or nerve cells connecting to the other parts of the brain and body. The cerebral cortex consists of four lobes (or zones) that regulate different body functions (Figure 1). More detail about each lobe is follows.

2.1.1. Frontal lobe

The frontal lobe is at the front of each cerebral hemisphere. It is one of the four divisions of the cerebral cortex. It regulates decision-making, problem solving, judgment, coordination, purposeful behaviour control, motor function, consciousness, and emotions.

The executive functions, which refer to the capacity to plan and carry out complex task, have close ties with frontal lobes [57]. This lobe has an important role in long-term memories that are not task-based. These memories linked with emotions originated from the limbic system. The emotions are adjusted to normally fit generally acceptable norms [100]. The frontal lobe is highly susceptible to injury.

2.1.2. Parietal lobes

The parietal lobes are located in the middle section of the brain, posterior to the frontal lobes and above the occipital lobes. They are associated with incorporating sensory information from different body parts, knowledge of numbers and their relations [72], and in the objects manipulation. Functions of parietal lobes also include processing sense of touch information [144].

Parietal lobes consist of two parts located on left and right hemispheres. The left parietal lobe takes part in symbolic functions in language and mathematics. The right hemisphere regulates maps understanding (e.g. spatial relationships – regulate how objects are positioned in space).

2.1.3. Occipital lobe

The occipital lobe is located at the back portion of the brain. The occipital lobe is the visual processing centre of the human brain [97]. This is because the primary visual cortex, which
collects and interprets visual information, is located in this lobe [115]. It has been known that damage to this lobe leads to difficulties in recognizing objects, colours, and words [39].

2.1.4. Temporal lobes
The temporal lobes are located at the sides of the brain. The temporal lobes are accountable for auditory processing and is also the location to the primary auditory cortex [146]. Consequently, temporal lobes are important for interpreting sounds and language that we hear. These lobes also contain the hippocampus that plays essential role in the formation of new memories [12]. Damages to the temporal lobes can cause problems of speech perception, memory, and language skills [101, 173].

2.2. The Anterior Cingulate Cortex
The anterior cingulate cortex (ACC) is a part of the limbic system in the brain. It is particularly involved in early learning and problem solving in which some effort is required to carry out the task [7]. Anatomically, it can be divided into ventral and dorsal components [37]. The ventral anterior cingulate cortex (vACC) is connected with the amygdala, anterior insula, and hypothalamus. It is involved in evaluating the importance of emotion and motivational information [37]. On the other hand, the dorsal anterior cingulate cortex (dACC) is positioned in the middle surface of the frontal lobe which relates to the reward-based decision making [143, 191] and task difficulty prediction [36]. It also monitors the conflicts between responses [34], error encountered, and differences between the expected and the actual result [10, 123].

As pointed out above, the ACC is activated when there is a confliction within the user or when there is an error made. One typical task that has been used widely to investigate this is the Eriksen flanker task [59]. In this task, there is an arrow pointing either to the left or right. Four arrows with two on each side flank this arrow. Consequently, there are 4 combinations categorizing into two groups: compatible (<<<<, >>>>) and incompatible (<<<<, >>>>) [34]. For each trial, users have to indicate the direction of the middle arrow as quickly as possible. Another common task that also induces conflict is the Stroop task [141]. This task involves naming the colour of a word, which can be congruent (the RED word has red colour) or incongruent (the RED word has blue colour). The conflict is triggered because of the confliction between users’ reading abilities and the ink colour of the word. Some variations of this task include Counting-Stroop (count number of interfering stimuli, e.g. “three” presented four times) and Emotional Counting Stroop (similar to the Counting Stroop test but has repeated emotional words such as “murder” during the interfering part) [189].
2.3. Measuring brain activity

2.3.1. Different measuring techniques:

a. Magnetic Resonance Imaging (MRI)

One of the common imaging techniques is Magnetic Resonance Imaging (MRI). MRI constructs high quality images of body structures (2- or 3-dimensional) using magnetic fields and radio waves surround the scanning objects [109]. This technique is non-invasive, painless, and does not require X-rays or radioactive tracers. An MRI machine usually has a tube shape surrounded by huge circular magnets, which create strong magnetic field resulting in the magnetic alignment of the protons in hydrogen atoms. This leads the nuclei to produce a faint signal due to the magnetic alignment (a rotating magnetic field) that can be detected by the scanner. Images of the scanning area are then reconstructed by processing this information on a computer.

One main advantage of MRI is that it can provide a detailed and high-resolution image of a body structure. It can also detect tiny changes in a body structure. In addition, it is especially useful in imaging the brain compared to other imaging techniques such as CT (computed tomography) or X-rays because it can provide images with good contrast between different soft tissues of the body. However, due to the effect of the magnet, most common metallic items are not allowed during the scan (for example, people with pacemakers, metal clips, or any computer devices). In addition, subjects usually lay down on an fMRI bed with limited space for movement. This may cause the claustrophobia effects (fear of confined places).

b. Functional Magnetic Resonance Imaging (fMRI)

Functional magnetic resonance imaging (fMRI) is a special type of MRI procedure that measures brain activity. This is based on the fact that neuronal activities and the cerebral blood flow are linked. In other words, the amount of oxygenated (oxy-Hb) and deoxygenated haemoglobin (deoxy-Hb) in the blood flow changes due to the neural activity. This combines with the paramagnetic properties of the oxy-Hb and deoxy-Hb allow fMRI to construct the images of changing blood flow in the brain [117]. The most common approach towards fMRI uses the Blood Oxygenation Level Dependent (BOLD) contrast [93]. BOLD fMRI allows us to measure the ratio of oxy-Hb to deoxy-Hb haemoglobin in the blood. This does not directly measure neuronal activity. However, it does measure the metabolic demands (or oxygen consumption) of active neurons. Images constructed from fMRI during performance of different
tasks reflect what parts of the brain are active and can reveal which (and how) brain structures are associated, being activated together during task.

Figure 2. The 3Tesla Siemens Magnetom Skyra MRI scanner at Clinical Imaging Research Centre Bristol.

fMRI has excellent spatial resolution compared to EEG and can record signals from all regions of the brain. It is also a non-invasive method and does not involve radiation, which is safe for majority of subjects. However, all restrictions on participants during MRI scanning also apply for fMRI.

c. Functional Near Infrared spectroscopy (fNIR)

Figure 3. fNIR 1100 Imager with 4 photo-emitters and 10 photo-detectors (Biopac system)

fNIR is one of the non-invasive functional neuroimaging methods, which use near infrared spectroscopy (NIRS) for functional imaging purpose. Using this method, the brain activity is measured through hemodynamic responses associated with it. In 1977, Jobsis [96] first propose the idea of using near infrared light (NIR) non-invasively to monitor the cortical tissues’ state. Using light in the range of 600 nm to 900 nm, the skin, tissue, and bone are mostly transparent, while oxy-Hb and deoxy-Hb absorbs lights stronger [168]. Basically, lights with different wavelengths will have different attenuation caused by oxy-Hb and deoxy-Hb. This can then be used to estimate the relative changes in haemoglobin concentration in the blood. Typically a
fNIR device has light emitters and detectors, which are placed on the users skull so that the emitted lights are reflected following elliptical pathways (Figure 3).

Similar to fMRI, fNIR uses the Blood Oxygenation Level Dependent (BOLD) contrast to measure the related changes in localized cerebral blood flow. These two techniques are sensitive to similar physiologic changes and are often comparable. fNIR has the advantages of low cost and portability over fMRI and MEG (as seen in Figure 3). On the other hand, because of the limitations in light emitter power, fNIRs cannot measure penetrate more than 4cm of the cortical activity. However, because of its advantages and characteristics, fNIR has been used to measure users’ concentration, engagement, or similar features related to cerebral blood flow (e.g. [87, 167]).

d. Magnetoencephalography (MEG)

Figure 4. A user is performing a task in the MEG machine1.

MEG is one of the functional neuroimaging techniques, which uses very sensitive magnetometers to record magnetic fields surrounding the head. This magnetic field is caused by the electrical activity within the brain. The first MEG was conducted in 1968 by Cohen in a shield room using an induction coil magnetometer [45]. The MEG measurements’ quality was significantly improved later by Cohen using SQUIDs (superconducting quantum interference devices) [46]. Nowadays, a MEG machine uses arrays of SQUIDs as its magnetometer. This allows capturing MEG efficiently and rapidly for most of the head.

Because of its characteristics, MEG is usually used in functional brain research investigating perceptual and cognitive brain processes, and neurofeedback. Main advantages of MEG are

---

non-invasive and non-ionizing radiation, independence of head geometry, and can be used to complement other brain activity sensing techniques (such as EEG, fMRI, and PET). In addition, because magnetic fields are less distorted than electric fields by the skull and scalp, MEG can produce better image at better resolution compared to an EEG. However, MEG is sensitive to ferromagnetic items so similar restrictions on which items are allowed in the fMRI scan room also apply to MEG.

2.3.2. Electroencephalography (EEG)

Electroencephalography (EEG) is one of the brain imaging techniques that reads the electrical activity along the scalp generated by neuronal activity. The first EEG recordings were conducted in 1924 by Hans Berger, a German neurologist, using a radio equipment to intensify the electrical activity captured on the human scalp [24]. The principle of EEG lies in the fact that when neurons are activated, ions such as Na\(^+\), K\(^+\), and Cl\(^-\) are pushed through the neuron membranes [16]. A large collection of active neurons can produce a recordable electrical activity on the surface of the head. Consequently, this can be detected and captured by the EEG electrodes. However, this is still a weak signal and usually is amplified to be displayed on paper or stored on a computer [179].

*Figure 5. The first human recorded electroencephalography [139].*

EEG is usually used to analyse epilepsy by observing abnormalities in EEG readings [5], sleep disorders, and patients in coma. Although EEG has a limited spatial resolution compared to fMRI, it is a favoured as a tool for conducting research and diagnosis because of its high temporal resolution and portability.

EEG can be either invasive or non-invasive. Invasive EEG requires electrodes to be surgically implanted on the surface or within the depth of the brain. Non-invasive EEG only reads electrical activity from the head surface. Non-invasive EEG has benefits of repeating measurements to normal adults and children with virtually no risk or limitation [138]. Consequently, non-invasive EEG has been more favourable in developing applications for normal and healthy users.


**a. Electrode Positions**

As a convention, EEG electrodes are placed around the human scalp following the 10-20 International system (Figure 6) [95]. It is a conventional method and internationally recognized, to describe the scalp electrodes’ locations in an EEG experiment or test. This was introduced as a conventional standard so that EEG studies can be compared with each other and over time. This system was constructed based on the relationship of electrode locations and cerebral cortex. At first, there are two positions on the head used for positioning the EEG electrodes as anatomical landmarks: the nasion (between the eyes) and the inion (the lowest point at the back of the skull), as shown in Figure 6, left. Then the number 10 and 20 referring to 10% and 20% of the actual distance between adjacent electrodes compared to the front-back or right-left distance of the human skull (Figure 6, middle and right). The letters and numbers are explained as in Table 1 [121].

![Figure 6] Locations of the 10-20 system, as standardized by the American Electroencephalographic Society [121].

**Table 1. 10-20 international system acronyms explanation**

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>Frontal lobe</td>
</tr>
<tr>
<td>T</td>
<td>Temporal lobe</td>
</tr>
<tr>
<td>C</td>
<td>Central lobe</td>
</tr>
<tr>
<td>P</td>
<td>Parietal lobe</td>
</tr>
<tr>
<td>O</td>
<td>Occipital lobe</td>
</tr>
<tr>
<td>Even numbers</td>
<td>Electrode positions on the right hemisphere (e.g. 2, 4, 6, 8)</td>
</tr>
<tr>
<td>Odd numbers</td>
<td>Electrode positions on the left hemisphere (e.g. 1, 3, 5, 7)</td>
</tr>
<tr>
<td>A, Pg, Fg</td>
<td>Earlobes, nasopharyngeal, and frontal polar site respectively.</td>
</tr>
</tbody>
</table>
b. Frequency ranges

The captured EEG recordings are usually presented in rhythmical form. Using a Fourier transform, the captured signals can be separated into various rhythms of different frequencies that can be considered independently. This is useful in comparing, estimating, and classifying users’ mental states and on-going tasks. While there is no accurate agreement on the min and max frequencies for each type, we can define five main types of EEG waves [31] as shown in Figure 7.

![Delta, Theta, Alpha, Beta waves](image)

**Figure 7. Different type of EEG frequency bands examples [121]**

- **Delta** (less than 4 Hz): This frequency is dominant in EEG of up to one-year infants. It also appears in the last two stages of sleep (NREM stage 3 and REM).

- **Theta** (4 to 7 Hz): This band is dominant in EEG of users with drowsiness. Theta waves may also appear during hypnagogic states such as hypnosis, deep daydreams, and the state just before falling asleep (preconscious stage).

- **Alpha** (8 to 13 Hz): This band has strong appearance when users are relaxed. It is also connected to variations in arousal or sleep [161]. Usually the eyes close leads to the rise of alpha rhythms. Additionally, alpha frequencies reduce when users are drowsy and open eyes. A variant of the alpha band named Mu-rhythm is occasionally captured in the motor cortex and reduces its amplitude with the intention to make a movement.

- **Beta** (14 to 30 Hz). This band is considered as an index of cortical arousal [153]. Beta frequencies with low amplitude are usually connected with active or nervous thoughts and active concentration.
• **Gamma** (above 30 Hz, up to 100 Hz). Gamma frequencies may be associated with problem solving, anxiety, and awareness. This band also increases its appearance during perception and memory formation as well as associative learning [131]

c. **Artifacts**

Because the sources of EEG signals are generated by activities of the brain’s neurons, the EEG signals are usually weak and need strong amplification [24]. Consequently, EEG signals can be easily contaminated by other sources of signals with equal or much higher amplitudes. These artifacts originate from non-physiological as well as physiological sources. Therefore, it is assumed that EEG signals are usually buried in noise from various sources [180, 194]. These noises are summarised and categorized as:

- **Eye (EOG) artifacts**: originate from movements of eyeballs, ocular muscles, and eyelids [11]. This type of noise is caused by the impending differences between the cornea and retina, unrelated to light simulation [121]. The amplitude of the signal can be quite large compared to the cerebral cortex. Techniques to counterpart eye artifacts include using EOG electrodes as references or using signal-processing methods (e.g. [25, 98, 136, 193]).

- **Electrocardiogram (EKG or ECG) artifacts**: are generated by the heart and can be mistaken for spike activity or cause rhythmic activity EEG. ECG artifacts often change amplitude accordingly to user’s breathing because breathing cause changes in the heart’s position relatively to the head [69]. To counterpart this, reference recordings of both ears can be used to produce less heart artifacts.

- **Electromyography (EMG) artifacts**: caused by the movements of the user’s head, jaw, tongue, or body. These artifacts may increase during tasks which include movement of the facial muscles [44]. To reduce this, participants of EEG experiments are usually advised to sit still and minimize unnecessary body movements. If unavoidable, EMG artifacts can be removed manually or automatically using signal processing after capturing in a similar way to EOG artifacts [133].

- **External artifacts**: originate from outside the human body such as power-line noise (50-60Hz depend on geological region), changes in electrode impedances, interferences by electrical generators, etc. These artifacts can be prevented by proper shielding and EEG signal filtering (e.g. a notch filter to remove power-line noise) [66, 138]
2.3.3. Active and passive Brain Computer Interface

Traditionally, EEG applications have focused strongly on providing a new communication or control channels for severely disabled people [192]. However, because of the continuous enhancements in the reliability and usability of BCI systems using EEG, there is an increasing interest in developing EEG based BCI applications for healthy users [199]. Although the EEG may be used in the same manner of continuous control and communications, it can be extended to assist and measure the inner states of users and also takes more passive way. Because of this, a new classification of BCI systems has been proposed, categorizing BCI applications in general, specifically using EEG into active, reactive, and passive BCI systems [137, 198-200].

- **Active**: an active BCI system detects signals triggered by the input of an activity on the brain. The signals are used for controlling an application directly and continuously by the user. Some examples of this are basket paradigm [106], the Hex-O-Spell [33], and imaginary movements [48].

- **Reactive**: a reactive BCI system detects brain signals resulting from external stimuli. The signals can be used to control an external application but not directly modulated by the user. An example of this is a P300 speller [108] where the user silently count the number of flashes over the intended target. Another example is the Steady State Visual Evoked Potential (SSVEP) where a looked object flashes at specific frequency triggering an unique brain response [111].

- **Passive**: a passive BCI system detects brain activity activated without voluntary control purpose. This type of BCI is mainly to enrich the human computer interaction with hidden information about the current underlying user states. An example of this is a system detecting Error Related Negativity, a brain pattern triggered when the user makes a mistake or gets confused about the last made action [184].

All of the above three BCI types have been investigated widely, focusing how to detect the appropriate brain signals to build a corresponding controlling and communicating system. There are not many studies focusing on guidelines and how to harness the types of signals best to benefit healthy users.

2.4. Interpreting brain activities in emotions and cognitive states

Currently, there are several low cost EEG headsets available that output users’ emotions based on measurements of their brain activity. For example, NeuroSky MindWave outputs Attention
and Meditation. Another example is the Emotiv EPOC which outputs excitement, engagement, boredom, and meditation. However, it is currently unknown how these emotions are calculated by these devices. Therefore, these devices are considered as black boxes, which make it harder to replicate results of any research with a different headset. On the other hand, these EEG devices can allow researchers access to raw EEG signals in real-time. This feature allows researchers to calculate the emotions or cognitive states from the raw EEG as from other devices. This section overviews current result related to measuring emotions using EEG raw signals.

2.4.1. Meditation
Recent studies have illustrated some insights into the neural basis of the meditation techniques by investigating their EEG reflections. Travis & Shear [178] summarise different types of meditation associating with their own EEG characteristics. For example, focused attention meditation (intended control of attention and mental processes) is associated with gamma and high beta band (20-30Hz) [120]. Open monitoring meditation such as Zen meditation, Sahaja yoga is associated with theta frequency band [20, 134]. Automatic self-transcending meditation technique has dominant low alpha frequencies (8-10Hz) [53, 177]. Additionally, significant increase of theta across all brain regions and alpha in the posterior region of the brain were found while participants were in meditation condition [110]. Given these results, off-the-shelf EEG headsets with access raw EEG signals can calculate the meditation values in real-time in order to benefit computer users accordingly.

2.4.2. Task engagement
According to Matthews et al. [125], task engagement can be defined as the effortful concentration and striving towards task goals where task demands and personal characteristics may influence this pattern of processing. In another work, Matthews et al. noted that effort investments in the task would increase task engagement [126]. Studies have shown a positive correlation between EEG engagement and task demands including stimulus complexity processing and the requirement for attentional resources allocation [25, 171]. EEG engagement correlated with task loads in not only simple vigilance and memory tasks but also in complex simulation tasks such as in radar operations simulation environment [25, 26]. Consequently, if an application requires low level of task engagement in using, it also requires low level of task demands, task loads, and workloads. In other words, it can be considered easier to use when compared with applications requiring higher task engagement.
Pope et al. [149] present a measurement of task engagement from EEG as $beta/(alpha + theta)$. Further studies [71, 150], with the same measure were used to show the performance quality benefits and reduced mental workload. These positive findings have been also replicated using extended periods of task performance [70] and a vigilance task [129].

2.4.3. Evaluate interactive systems using different types of emotions

Traditionally interactive programs are evaluated by investigating performance (i.e. [160]), or users’ behaviours. However, with programs designed for using with ease and pleasure, an evaluation method measuring users’ affective or inner states is more suitable. Usually this is done by questionnaires answered by participants during the experiment. However, they occur after the event when important issues may be forgotten. Additionally participants may not be aware of their states or might simply guess. Neural signals, measured from the brain can better reflect a users’ current state and provide an evaluation metric. Consequently, evaluation methods using neural signals can provide task designers a direct and intuitive way to evaluate their interactive applications through the interpretations of users’ brain signals.

2.5. Event-related potentials (ERP)

Event-related potentials (ERPs) are signals generated in the brain by groups of neurons in response to specific events or stimuli (either perceptual, cognitive or motor event) [30]. This is different from the brain activity generated in response to volunteer self-paced tasks. These EEG potentials, which are time-locked to the triggered events, can be seen as a specific kind of EEG evoked potentials such as Visual Evoked Potentials (VEP) or Auditory Evoked Potentials (AEP). ERPs reflects the underlying combined activities of the postsynaptic potentials generated when a large group of neurons fire in synchrony while processing information [145]. Consequently, captured EEG in related areas resulted from the neurons fire shows modifications called potentials. However, because the EEG containing the ERP is usually contaminated by various sources of noise, time-locked ERPs are often averaged to remove the random noise while keeping the ERP components. Up until now, there have been various ERPs discovered related to different types of stimuli. However, researches on ERPs also extend on how to harness the ERPs. In order to do that, the first challenge is to detect ERP not only in averaging many single trials, but also to detect it immediately at the single trial level. In the following sections, we will cover some ERPs that have been utilised in assisting HCI users.
2.5.1. P300 evoked potentials

P300 is a late ERP component triggered in the process of decision-making and can be considered as an index of attention and memory [113]. Its amplitude correlates with lower probability and higher discriminability of targets where its latency increases when it is harder to discriminate targets from standard stimuli [116]. P300 is evoked in an “oddball” task where low-probability targets and high-probability non-target (or "standard") items are mixed in a random appearance, and with an inter-stimulus interval (ISI) of between 1s and 2s [188]. It can be considered as an autogenous potential, since its occurrence relates more to the person’s reaction to the stimuli rather than the physical attributes of the stimuli itself [56]. In addition, P300 patterns can be auditory, visual, or somatosensory ERP [29].

P300 has many desirable qualities that benefit its implementation in interactive applications. For example, a P300 pattern can be evoked in almost all users with little change in measurement and design technique. This results in a simplify user interface design and enables greater usability. In addition, it is consistently detectable in response to precise stimuli [56]. However, the detection of P300 usually depends on averaging multiple trials to extract the signals. Consequently, the speed of a P300-based interface is determined by the speed of averaging process and post-recording signal processing [54]. In HCI, various attempts have been made to utilize P300 in designing interactive tasks. For example, Farwell and Donchin [65] provides a P300 speller, consisted of a 6x6 grid of characters, to assist the user to “type” at the speed of 3.4-4.3 chars/min. Other examples include selecting a physical object on a tabletop [197] or call a contact’s phone number on a mobile phone [38].

2.5.2. Error-related negativity (ERN)

ERN is an error-related potential on error trials, which is also a type of ERP. It was first reported in 1990 by Gehring et al. [74] and Falkenstein et al. [62]. An ERP is triggered in a user’s EEG when the user has committed a mistake in a multiple choice reaction time task [28, 63, 75, 163]. It also elicited, with lower amplitude, when the user is confused about the last action [47]. Although the ERN waveform has varied amplitude across experimental condition, its latency seems to be very consistent [89]. ERN is reportedly elicited in fast reaction time tasks within 150ms following the awareness of an erroneous response [40, 64, 76, 156, 159, 187].

a. Source localization

Recent studies into source localization of ERN have suggested that ERN is generated from the medial frontal cortex, particularly the caudal region of the Anterior Cingulate Cortex (ACC) [47,
51, 183, 195], as shown in Figure 8. This conclusion is also supported by fMRI studies [90, 94], brain lesion research [172], and dipole source modelling [51]. In addition, there are different theories of how ERN are generated within the brain, whether in response to or detection of errors [28], by monitoring confliction [35], or by continuous comparisons [64]. However, it was pointed out that ERN has frontal-central distribution [47, 196] that shows which locations (in the 10-20 international system) that are more likely to be able to detect ERN patterns.

Figure 8. Location of the anterior cingulate cortex (ACC – in red) in the human brain.

b. Eriksen flanker task – Multiple choice reaction time task

Typically ERN patterns are elicited when a subject committed an erroneous response or confused about his last action in a multiple choice reaction time task. In this task, subjects respond to a stimulus, auditory or visual, as quickly as possible. One common multiple choice reaction time task, which has been largely studied to trigger ERN, is the flanker task [59]. Procedure of each trial in this task is follow. At first, participants were presented by a target letter (either S or H), which was flanked by 2 distractors on each side. Combinations of target letters and distractors formed compatible stimuli (SSSSS and HHHHH) and incompatible stimuli (SSHSS and HHSBH). Participants had to produce a right- or left-hand action in response to the just presented target letter. Once the participants committed errors, ERN patterns appeared within 150ms and had the frontal-central distribution [63, 75]. Another version of Flanker task is using arrowheads (< and >) instead of letters [142, 182, 195]. Consequently, the two compatible- and two incompatible stimuli are <<<<, >>>>, <<<<, and >>>> respectively.
c. Observation tasks

To date most examples of ERN have focused on detecting the signal by recording signals from the executer of the task. Recent endeavours [50, 132, 182] suggest that ERN signals even appear while observing another user making errors. These studies confirm the existence of negative potential within 250ms after the event onset and analysis of the origin of the signal confirm that they are ERN signals. In these studies the observer did not have much opportunity to anticipate the actions of the executer and often relied on the answer displayed. In HCI scenarios, the executer’s actions may become visible to the observer even before the action is committed. For example, in collaborative tabletops, an observer has awareness of an executer’s actions through their reaching gesture [147].

To date, all studies of ERN in observation tasks only look at situations where the outcome of the action becomes visible around the moment where the action is committed. Furthermore, these studies focus on detecting the ERN using averaging methods for the purpose of confirming the pattern’s existence. We are not aware of studies that investigate the context where the observer can anticipate the executer’s actions through their gestures towards the target.

d. Co-operative vs. competitive tasks

Recent endeavours [22, 50, 132, 182] suggest that ERN signals even appear while observing another user making errors. For example, Miltner et al. [132] reported the generation of an ERN potential when participants observed errors committed in a choice reaction time (RT) task. Moreover, van Schie et al. [182] reported an ERN pattern elicited in an observer when observing another person performing a modified Erikson flanker task [60]. Other researches have confirmed these results by showing observed ERN following observations of other’s errors [22, 132]. These results suggested that similar neural processes trigger the detection of a person’s own error as well as the detection of others. Interestingly, de Bruijn et al. [50] aimed to disentangle the dependency of ERN on error or reward. Their results show that the performance monitoring, reflected by the ERN, is not directly dependent on reward but is error specific. As a result, our studies will focus on the error awareness of the observers.

e. Different attempts to classify ERN

In an interactive application, ERN will be most useful when it can be detected immediately after it happens. This requires the online detection of the ERN pattern (as opposed to offline) from EEG signals. One attempt to detect ERN patterns online is by caching and averaging a small number of patterns in a limited time in order to refine a reasonably clear ERN [166]. This
method requires a waiting time to collect sufficient signals and therefore delays the progress of error correction as it goes through multiple trials.

Various attempts have also been made to detect ERN from a single trial. One example is from Ferrez and Millan [68] where they train a Gaussian classifier to recognize error and correct trials on a single trial basis but after the experiment was completed (so considered an offline single-trial approach). Another example is from Dal Seno et al. [49] who used a Linear Discriminant Analysis (LDA) classifier to cross check with every spelled letter using a P300 speller. The reported performance of the online version varies from 58% to 69%. This result is just better than chance but still encouraging, as this is one of the few attempts at detecting ERN online and from a single trial.

All of the above experiments have been carried out on EEG headsets with up to 256 electrodes and with sampling frequency up to 16kHz. Most importantly these approaches rely on having access to Fz, Cz, Pz and Oz channels in the 10-20 international system. We are not aware of any attempt to detect ERN from a low-cost portable headset where the number of channels and sampling frequency are limited. Hence, we investigate if a low cost, features limited EEG headset can detect ERN and benefit to HCI.

2.6. Data acquisition and processing

2.6.1. Hardware

Traditionally most ERN experiments are performed using expensive, non-portable sensing devices such as NeuroScan systems from NeuroScan™, g.BCI sys from g.tec™ or ActiveTwo from Biosemi™. These devices have the benefits of not only more sensing but also higher and user selectable sampling frequencies. These systems also introduce less noise in the signals because of better sensing electrodes and integrated amplifier/ converter module.

In a comparison off-the-shelf EEG headsets (e.g. Emotiv EPOC, NeuroSky MindWave, etc.) have at best 14 channels, a sampling frequency of about 1 kHz and introduce more noise due to wireless transfer of EEG signals. Due to these limitations it is not clear if these devices can capture ERN patterns. However, due to their price and portability their use is being increasingly explored within HCI.

Eleven studies using EEG devices are presented in this thesis. Ten of them were with Emotiv EEG neuroheadset (where one of them was placebo); one study was with EbNeuro BE-MRI
system. The used EEG devices belong to two different categories as discussed above: Emotiv EEG neuroheadset is a low cost, off-the-shelf, commodity headset whereas EbNeuro BeLight EEG system is a clinical, high quality and expensive EEG system. The specifications and capabilities of each device are described as follow.

a. **Emotiv headset**

This device is developed by Emotiv™ and first introduced in 2003. This device is favoured by many HCI researchers because of its low-cost and portability while still offering a feature to capture raw EEG signals.

Emotiv EEG headset has 16 channels in the 10-20 international system. Two of them are used for references at P3/P4 as Common Mode Sense (CMS) active electrode/ Driven Right Leg (DRL) passive electrode. EEG signals are collected at 2048Hz internally and then down-sampling to 128Hz before sending to the output. The output signals are from 0.2Hz – 45Hz and digital notch filtered at 50Hz and 60Hz.

Emotiv offers portability by using a Bluetooth wireless at 2.4Hz band. This device is customized to general users and gamers by using saline solution instead of neurogel and long-hours of internal battery life (about 12 hours). It has been reported that users can comfortably wear Emotiv for at least 1 hour of continuous use [58]. See Appendix 1 for more technical information.

b. **BE-MRI system**

This system is developed by EbNeuro. It is specially designed to work in both EEG and MRI environments where all electromagnetic signals from outside the signals box must be shielded. This provides low-noise signals at the signal capturing level.

The system includes a BE-Light amplifier, which has sampling frequency up to 8 kHz/channel. This module includes many advanced features such as: high signal quality, artefact rejection, and electrode montage impedance check directly on the head-box. This device can detect all channels in the 10-20 international system. However, this feature depends on the cap connected to the Amplifier module.

The setup of a study using BE-MRI requires longer cap preparation and need conductive gels/pastes. Besides the high price, this makes it more difficult for the device to be widely used for general public and HCI researchers.
2.6.2. Evaluate classifiers using the Receiving Operating Characteristic

A receiver operating characteristics (ROC) is a technique to visualize, organizing, and evaluate a performance of a classifier(s) in a graph [67]. This technique takes inputs are the detection rates (or true positive rate – TPR) and false alarm rates (or false positive rate – FPR) to produce the trade-off between them [91]. It is based on statistical decision theory and was developed to solve problems of signal detection in the early 1950s [128]. Additionally, ROC graphs have been increasingly used in different area such as visualizing and analysing behaviours of diagnostic systems [175].

![ROC Graph](image)

*Figure 9. An example of ROC graph*

Since one of the first use of ROC in machine learning presented by Spackman [170] to evaluate algorithms, it has been used widely in the machine learning community to compare different machine learning techniques. This is because the accuracy of a classifier is not good enough to prove good performance [67, 151]. By adjusting the TPR and FPR, the specific curve of the classifier is plotted on the graph of those two axes (as seen in Figure 9).

When comparing different classifiers or algorithms using ROC graphs, a normalized unit called the area under the curve (AUC) is used [67]. An AUC has range of [0, 1] where 0.5 represents a classifier with detection rate equal to chance and 1 represents a perfect classifier with 100% detection rates. Therefore, by calculating the AUC value of a classifier we can know how effective that classifier is in detecting the required object.

2.7. Summary

In this chapter we have covered three related topics. First, the brain structures are covered anatomically and functionally. It is important to understand how the brain is formed and how
different parts of the brain function. This potentially helps us to determine which locations of the brain to look at in the captured signals.

We also explored different techniques to measure brain signals. They are MRI, fMRI, MEG, fNIR, PET and EEG. Among them, EEG was favoured in this thesis because of its low cost and portability (compared to MRI, fMRI, MEG, PET), and its high temporal resolution in real time (compare to fMRI, fNIR). Besides, research about EEG has been widely investigated for disabled people for a long time but there are limited explorations about interaction designs employing EEG. This leaves a wide research area that is interesting and useful for computer users.

In addition, we covered in more details about EEG as it is the chosen sensing technique in this thesis. The continuous usage and ERP are reviewed. In the later chapters of this thesis, we will cover two types of EEG-based system: active and passive systems. For each scenario, we will demonstrate and provide guidelines of how to design interactive applications using EEG, particularly using task engagement and Error-Related Negativity.

In order to do that, we will firstly investigate how to design an active BCI-based system where task engagement, calculated from users’ neuronal signals, can be used as an input control. This imposes two separate usage of task engagement: (1) whether it can replace the traditional input controls; or (2) whether it is useful as an additional input beside the existing and traditional input methods. We will investigate this in the next chapter.
Chapter 3. Quantifying Task Engagement for use in Gaming Applications

We started our investigations aiming to design an active BCI-based system where neuronal signals captured from users can be used directly as an input (e.g. in gaming applications). Therefore, this chapter aims to provide fundamental understanding of the speed and accuracy with which users can control virtual objects using only their task engagement levels (mental effort activity), of their brain. Therefore we present a series of studies to investigate the use of EEG techniques in the context of gaming. We first investigated the accuracy with which users could select a single target from multiple visual states. We found that users able to move quickly and effectively in both directions (increasing engagement/decreasing engagement) for up to five discrete states. In a second study, we found that users take a linearly increasing amount of time to increase and decrease their level of engagement. Finally, we investigated the practicalities of simultaneously using a traditional input device and engagement in a gaming scenario. Experienced users were more accurate in this parallel input task than novices. Based on these experimental results we discuss several engagement-based game design principles.

3.1. Potentials of commodity EEG headsets

Electroencephalography (EEG) kits (such as the Emotiv EEG and Neurosky headsets) are readily available off-the-shelf, allowing designers to explore novel Brain-Computer Interfaces (BCI). Applications using the EEG signals captured through non-invasive electrodes range from gaming [88] to robot control [8]. Researchers have mainly demonstrated the capability of these controllers through specific point-designs in applications such as user task classification [113] to controlling a wheelchair [114].

In these applications, EEG signals are commonly used to capture the affective state (e.g. relaxation [88]) of the user which can then be mapped to controlling virtual [114] or physical objects [88]. Researchers have explored a range of possibilities with sensed parameters, ranging from arousal [52] to engagement [27] and relaxation [88]. However, there exist only a limited number of generic design principles that can guide application developers in mapping a sensed
parameter, such as the level of engagement, to virtual control. Answers to some of the fundamental control issues such as “how many discrete levels of engagement can a user comfortably control?” or “is it easier for users to control a parameter by becoming more engaged than less engaged?” are largely unexplored. Answers to such questions allow designers to have more generic forms of mappings of EEG signals to virtual control.

![Figure 10. A user wearing the Emotiv EEG headset in the final study setup that combines engagement control and traditional input.](image)

This chapter is targeted at producing generic design principles for controlling EEG sensed parameters based on task engagement (hereafter “engagement”). According to Matthews et al. [125], engagement can be defined as the effortful concentration and striving towards task goals where task demands and personal characteristics may influence this pattern of processing. In another work, Matthews et al. noted that effort investments in the task would increase task engagement [126]. Studies have shown a positive correlation between EEG engagement and task demands including stimulus complexity processing and the requirement for attentional resources allocation [25, 171]. EEG engagement correlated with task loads in not only simple vigilance and memory tasks but also in complex simulation tasks such as in radar operations simulation environment [25, 26]. Consequently, if an application requires low level of engagement in using, it also requires low level of task demands, task loads, and workloads.

Commercial EEG kits (such as the Emotiv) can sense levels of engagement in real time giving interaction designers the ability to exploit this fundamental mental state as a control mechanism for various interactive activities. Previous research has derived methods to successfully measure engagement using EEG [42, 148], but the low-level questions for interactive control listed above remain unanswered.
We implemented three experiments to generate principles for one-dimensional selection in the absence of environmental or task distractions, and a final ecologically valid experiment for controlling gaming parameters through levels of engagement.

From these experiments we observed that (a) users can comfortably control five discrete levels of cursor movement using engagement; (b) the time taken to move across levels of engagement is proportional to the distance required to displace a cursor; (c) that engagement affords bi-directional control (i.e. moving to higher and lower levels require similar amount of effort); and (d) in multi-tasking activities, engagement is a useful sensed parameter, but is mastered through repetitive task exposure.

The main contributions of this chapter are: 1) a systematic exploration of parameters that influence mappings of engagement to virtual cursor control; 2) quantifying the number of discrete levels that can be controlled via engagement states from EEG input; 3) a series of guidelines for designers; 4) a demonstration of the practical value of our mappings in a gaming application.

3.2. Related Work

A natural use of EEG sensing is for controlling interactive devices. This has led to a number of interesting applications, including new input methods for people with severe motor disabilities [65], capturing emotional and cognitive user states during computer use [104], facilitating interface adaptation based on the user’s cognitive states [79], and augmenting traditional input controls with an additional emotional state, to enrich game experiences [1].

Computer games have progressed from keyboard and joystick input to rich physical movements (such as the Nintendo Wii and Microsoft Kinect). Future games are likely to take advantage of the BCI to increase user involvement. One discovery that has made BCI more interesting for gaming has been to use imaginary movements to produce similar brain signals as real movements [78]; it is then possible to create a BCI based agent which replicates the user’s imagination [114].

The incorporation of BCI input technology into games can be categorized into three different groups:

- **Traditional games with a BCI** were originally designed for mouse and keyboard input but have been adapted to utilize BCI as the only input control for the game.
Examples of these games include the walking game [118], or playing Ping-Pong using a BCI controller [3].

- **Non-traditional games with a BCI** afford the design of conceptually different games to those possible with traditional input mechanisms. For example, these games may utilize BCI capabilities to help players improve their mental awareness. As an example, in Brainball, gamers must relax to win the game [88].

- **BCI assisted games** augment BCI input with traditional controls to provide a richer experience. For example, if the user is frustrated when attempting to complete a task, the game could assist the player, or if their level of engagement is not adequate, the game difficulty could be elevated [77]. Mindflex (mindflexgames.com) is another commercial game that requires users to guide a ball through an obstacle course. The height of the ball is controlled by excitement while the movement in other directions is controlled by a knob.

One promising and highly sensitive parameter that is of interest is engagement. Pope et al. [149] present a measurement of task engagement from EEG as $\frac{\text{beta}}{\text{alpha + theta}}$ at channels Cz, Pz, P3, and P4. Further studies [71, 150], with the same measure were used to show the performance quality benefits and reduced mental workload. These positive findings have been also replicated using extended periods of task performance [70] and a vigilance task [129]. Interestingly, Fairclough et al. [15] raised concerns about directly applying engagement as a physiological input controller. In particular, Fairclough et al. questioned whether levels of engagement were sufficiently discriminable for input control.

In this chapter we further study the discriminability of discrete levels of engagement for input control. Our study aims to provide design guidelines for the above listed styles of games. Additionally, the third experiment in this chapter uses a BCI assisted gaming style to demonstrate and evaluate the ecologically valid setting and the usability of our guidelines in a gaming environment.

### 3.3. Factors Influencing the Design of a System using Engagement for Control

There are several factors that influence the design of EEG based systems. This section identifies four such factors: engagement discretization, selection technique, visual feedback, and engagement calibration.
3.3.1. Engagement Discretization

EEG devices for measuring engagement provide a stream of constantly changing brain activity. Engagement can be measured as the ratio of the change in power-spectral density over a certain time window [148]. Depending on the length of the window, one can measure long-term or short-term engagement. For example, the Emotiv EPOC neuroheadset measures short-term engagement in a short-time (overlapping) window with a step of 0.25sec and reports a number from zero to one. Continuous input control makes immediate use of this value. For example, continuous inputs can enhance-first person shooter games: if a player is highly engaged, small movements that make targeting easier are executed more often.

However, such measures of engagement often spike and change quickly, making on-screen parameter control difficult. To provide smoother and accurate control, the input range can be discretized into a smaller number of values or levels. This has the effect of reducing the accuracy required by users to select a specific value along a parameter range. Discretization can benefit game design, for example, by allowing the designer to alter the number of discrete levels for game balancing. There are no reported results describing how to suitably discretize this space; this is one of the central contributions of this chapter.

3.3.2. Mapping Method

The aim of a discrete mapping is to translate a level of human engagement into one discrete command. There are at least two types of common mappings that a designer could choose from: absolute and relative mapping.

a. Absolute mapping

Absolute mapping updates the position of the cursor regardless of the user’s previous state. This can be done either linearly or non-linearly. In the former, all of the engagement states are given the same weight whereas in the latter different states afford different weights. In linear mapping, the position of the cursor is a direct one-to-one function of the current value of an emotion. Commercial games such as MindBall (www.mindball.se) and Mindflex (mindflexgames.com) use absolute mapping from relaxation level to the distance moved by a physical ball.

b. Relative mapping

Relative mapping positions the cursor based on the difference between the current and the previous engagement states of the user. This can be performed in different ways. One approach is to calculate the distance between two consecutive engagement states and with a function
(linear or non-linear) calculate the displacement that should be applied to the cursor (e.g. this is based on the operation of the computer mouse). Alternatively, the cursor displacement can be calculated relative to a neutral state, instead of a previous engagement state.

3.3.3. Selection Technique

A selection technique allows users to pick a state after moving the BCI-based cursor into the required level. There are two possible ways for selection to occur on BCI systems: (1) within band, i.e. using BCI capabilities and (2) out-of-band, i.e. using a traditional input controller. Within-band capabilities include dwelling, hovering upon one state for a length of time, and changing the level of an emotion for triggering selection. Out-of-band can occur via a mouse double-click, a button press, or voice recognition. Our study employs a combination of both techniques.

3.3.4. Visual Feedback

This is one of the main building blocks of control systems and is especially important in Operant Conditioning based BCI systems. For example, it is a key component of P300 based systems. Buttons or objects flash over a period of time on screen and users are asked to count the number of times the interested object flashes to help them focus on the object and improve the P300 signal [194].

Different visual representations lead to different effects. For instance, Ramos et al. [152] showed in their study that users performed better with full visual feedback compared with spatial feedback. Many BCI-based outputs have been used for input control. However, there is little research that investigates how well a user can control these input states in different conditions and different forms of visual representations. We investigate the effect of visual representation on discrete state control.

3.3.5. Calibrating engagement

To utilize engagement as an input for a system, it is necessary to map it to a scale (i.e. 0 to 1) where the maximum and minimum values are known and used as boundaries of the scale. Although correlations exist between engagement and the power spectrum of EEG frequency bands [42, 148], there is limited knowledge on how to find these limits. Better known is how engagement reaches low values when users close their eyes and relax, and reaches high values when users perform a three choice vigilance task (3CVT) [27]. Therefore, one solution to the mapping problem is using the values of each user obtained from the eye-closed task and 3CVT.
as the initial minimum and maximum values (see next section). These values can be adaptively changed during the task later. We use this approach to ‘calibrate’ engagement in our studies.

3.4. Experiment Setup and Signal Processing

3.4.1. Experiment Setup

We set out to better understand the parameters of interaction using engagement as an input for control. To do this we ran a series of experiments to understand user capabilities. This section describes the experimental setup and signal processing required. The following sections describe individual experiments.

We used the Emotiv EEG neuroheadset (a readily available and relatively cheap EEG based controller) in our studies. This headset collects EEG signals from 14 channels in the 10-20 international system [138]. The SDK provides access to raw EEG signals in real time. It also interprets excitement, meditation, and engagement intensities and represents those values on a scale from 0.0 to 1.0. However, the methods of extraction are ‘black-box’ and not revealed to researchers and consumers. To ensure the generalizability of our results we calculated the values of engagement from the raw EEG signals using the methods suggested by Pope et al. [148].

To calibrate participants perform the eyes-closed and three choice vigilance tasks (3CVT) for 5 minutes each. These values are mapped on a scale from 0 to 1 for the later control task. Participants first sat comfortably and relaxed with their eyes closed, but without falling asleep. This relaxed state was to measure the lower bound of the engagement of a participant (relaxed wakefulness state [27]). Participants then performed a three choice vigilance task (Figure 11).

The three choice vigilance task integrates features of common neuropsychological tests of vigilance. It also includes simple or choice reaction time tests and continuous performance tests [6]. It includes a primary object (presented 70% of the time) and two secondary objects. EEG signals collected from participants performing this task were used to calculate the higher bound of the engagement of each participant.

A trial started with an object appearing on the screen for 200ms; participants were asked to press the spacebar using both index fingers as soon as possible within two seconds after a primary object was shown, and do nothing with secondary objects. A message, which showed if the user performed correctly or not, was displayed for two seconds. There was a five second pause before the next trial began.
After performing the eyes-closed and 3CVT tasks, participants began the cursor positioning tasks (described later) - the on-screen cursor position was moved by users modifying their state of engagement.

![Three objects of the 3CVT. Objects appeared without captions in the experiment.](image)

**Figure 11. Three objects of the 3CVT. Objects appeared without captions in the experiment.**

### 3.4.2. Signal Processing

We conducted a series of steps to process raw EEG signals (see Figure 12). Raw EEG signals from the eight channels F3, F4, FC5, FC6, P7, P8, O1, and O2 were collected from the Emotiv neuroheadset. These channels were chosen because they are the Emotiv's closest to those used by Pope et al. The signals then were filled into a two second window. The program only starts after the first window is completely filled. This window of signals was filtered through different band-pass filters to obtain a combined power of alpha, beta, and theta.

![Signal processing procedure](image)

**Figure 12. Signal processing procedure**

The Engagement Index (EI) representing instantaneous user’s engagement was calculated by the equation (1). This formula captures engagement best [42, 148]. The window continually shifted and produced a new instantaneous Engagement index corresponding to the Emotiv’s sampling rate (128Hz). The values outputted from the closed-eye and 3CVT tasks were then averaged
separately over all instantaneous engagement indices to measure lower and upper bounds of the participant. These two values were used for the third task.

For the third task, the engagement value (on \([0, 1]\) scale) was an average of 32 instantaneous Engagement indexes (32 2-second windows). As the Emotiv has a sampling frequency of 128Hz, this also means a new engagement value is captured every 250ms. This interval is equal to the output interval of the Emotiv Affective Suite. The EI was obtained as the same way as in the previous two tasks. However, the EI was mapped onto a linear scale by the formula:

\[
EV = \frac{EI - \text{min}(EI)}{\text{max}(EI) - \text{min}(EI)}
\]

Where:

- \(EV\) = engagement value on the scale \([0,1]\)
- \(EI\) = average Engagement Index over a 2sec window and shifted every 0.25sec
- \(\text{min}(EI)\) = minimum Engagement Index (initial value was from eye closed task)
- \(\text{max}(EI)\) = maximum Engagement Index (initial value was from 3CVT)

The maximum and minimum EI values were monitored and adapted for the user through an adaptive module that widened the engagement index band if a new EI was smaller than \(\text{min}(EI)\) or larger than \(\text{max}(EI)\). For example, if the detected EI reached a larger value than \(\text{max}(EI)\) for a period of at least 250ms. This detected EI was stored in an array of size \([1x20]\) where all EI equal or larger than \(\text{max}(EI)\) were stored. The averaged value of this array’s elements was used as the updated \(\text{max}(EI)\) if there was a change. Similar rule applies to the case of \(\text{min}(EI)\). This process was to cope with the case when \(\text{max}(EI)\) and \(\text{min}(EI)\) obtained from the first two tasks were not the largest and smallest EI values that the participant could obtain. The module detected and prevented noise which caused out-of-reach max and min indexes. The noise, which fell within the measured frequency bands, could be from body movement, facial actions and surrounding electrical equipment. It also smoothed the cursor movement by re-scaling the Engagement index when there was a sudden jump in EI. Although the \(\text{max}(EI)\) and \(\text{min}(EI)\) might change during the experiment, the ongoing trials would not be more difficult as users also learned how to reach these new levels (learning effect).

### 3.5. Experiment 1: Identifying levels of control

The primary goal of this experiment was to establish the accuracy with which a user could manipulate their engagement levels to control an on-screen cursor. A secondary goal was to examine the suitability of different visual feedback types.
This study was divided into two parts: 1A and 1B. The experimental setup remained identical between the two parts, the conditions and participants were modified, as reported at the beginning of experiment 1B.

3.5.1. Experimental Procedure and Interface

The Emotiv headset was placed on the user and the experiment began when the participant was comfortable. At first, each participant performed the eye-closed task and 3CVT to get their own min and max EIs. These values were used for the following cursor control task.

In this task, the START state is shaded light green, the END state is shaded bright green, and the current state is shaded yellow (see Figure 13). To prepare the participant, a three-second countdown begins. Once the countdown ends, the user moves the cursor by changing their engagement level. The mapping was as such: being less engaged moves the cursor left (in the linear mode) or towards the center (bulls-eye mode). More engagement moves the cursor rightwards or outwards.

![Image](image.png)

**Figure 13. Bull’s eye (left) and Linear (right) visualizations. The yellow circle (square) shows the user’s state, the light green circle (square) shows the start state, the bright green circle/square at the end state, and the red line/circle shows the cursor.**

The participants’ initial engagement level must match the start state before a trial begins. To achieve this, the participants must modify their engagement level to move the cursor into the start state, and keep it at a constant engagement for at least 200ms, at which point the trial begins. The users then adjust their engagement level to reach the target. Once over the target, the user presses a key on a wireless remote to indicate the end of the trial. Trials are terminated if the user requires more than one minute to move and select the target. In this study, we use a
one-to-one mapping of engagement level to cursor position. Participants were instructed to adjust their mental effort only. They were asked to not make body and facial movements (i.e. clenching, laughing) during the experiment. During the trials the experiment was continuously monitored to make sure participants followed our instructions.

The visual interface had a size of 1280×1024 pixels and was displayed on a 19-inch widescreen monitor. Participants sat 90cm from the screen and their line of sight was perpendicular with the display. Participants were then given ample time to practice the 3CVT and control mechanism with both of the visualizations. The cursor control task consisted of four blocks of 16 trials. There was three minutes rest between blocks. The study took on average 60 minutes per participant, including practice time.

3.5.2. Study 1A

The experiment used a 2×4×2 within-participants design, with factors:

- **Number of states**: 5, 10. Increasing the number of states requires users to more accurately make their selection.

- **Distance**: +0.29, +0.59, -0.59, -0.29. Engagement is measured on a linear scale from 0 to 1. Distance is a measurement along this scale - positive distance indicates movement from less to more engaged, negative movement is the opposite.

- **Visualization**: Linear, Bull’s-eye. A one-dimensional matrix layout or a bull’s-eye-like set of concentric circles (see Figure 13).

The factors visualization and number of states were counterbalanced between participants. The experiment consisted of four blocks (5B, 5L, 10B, 10L with B: Bull’s-eye, L: Linear) of 16 trials with the presentation of distances modified between blocks. Each block contained 4 repetitions of all distances. The order of blocks was randomly assigned for each participant.

**a. Participants**

Nine participants (6 male and 3 female) between the ages of 20 and 31 volunteered for the study. All were from a local university and all had heard of an EEG headset but none had used one before. All participants performed the eye-close task and 3CVT to obtain their minimum and maximum values of EI before performing the cursor control task.

**b. Results**

The total number of trials that were successfully completed was 552 out of 576. The average trial completion time for successful trials was 9.06sec (s.e. = 0.43sec). We carried out statistical
tests using a univariate ANOVA with Tukey post-hoc pair-wise comparisons to compare the effect of number of states and visualization.

There was a significant effect of the number of states on trial completion time ($t = -5.187, p < 0.001$), with five states significantly faster (6.93 sec) than ten states (11.32 sec). There was also a significant effect of number of states on false target selection ($t = -2.900, p < 0.05$): five states had an average of 0.6 false selections, ten states 2.4. Figure 14 (left) shows the mean task completion time per state for each visualization.

![Figure 14](image1.png)

**Figure 14.** Task completion time in sec (left) and average number of false targets selected (right) for each state (horizontal axis) and visualizations (B / L = Bull’s-eye / Linear Visualization) for Study 1A

We found no significant effect of visualization on trial completion time ($t = -0.618, p > 0.1$, linear = 8.79sec, bull’s-eye = 9.33s) or number of false targets selected ($t = 0.221, p > 0.1$, linear = 1.44, bull’s-eye = 1.33). An exit survey showed no differences in user perception of frustration ($p = 0.707$) or effort ($p = 0.928$) between the two visualizations.

![Figure 15](image2.png)

**Figure 15.** Task completion time in sec (left) and false targets selected (right) over distances for Study 1A

There is a significant effect of distance on trial completion time ($F_{3,25} = 13.490, p < 0.001$) and number of false targets selected ($F_{3,32} = 7.980, p < 0.001$). Post-hoc pair-wise comparisons showed users were significantly faster ($p < 0.05$) at decreasing their engagement than increasing it. There were significant differences in completion time within each pair: (-0.59, -0.29), (-0.59, 0.59), (-0.29, 0.59), (0.29, 0.59) (see Figure 15, left). However, there were significantly more
false targets when going to 0.59 than -0.29 and -0.59. None of the other pairs showed significance.

3.5.3. Study 1B

In Study 1A, we observed that users were significantly faster when selecting from five states over ten and that there were less false selections using less states. Based on this result, we predicted that users should be even faster when selecting from three states over five. Study 1B tested this hypothesis. It was run with an identical setup as study 1A, with the number of states and participants differing as documented below.

a. Setup Differences

We modified the levels of the number of states factor to three and five. An additional eight participants (5 male and 3 female) between the ages of 21 and 30 were recruited for study 1B. As per experiment 1A, all had heard of an EEG headset but none of had used one before. All participants performed the eye-close task and 3CVT to obtain their maximum and minimum values of EI before performing the cursor control task.

b. Results

The total number of trials that were successfully completed was 501 out of 512. All false targets selected were with 5 states. The average trial completion time for successful trials was 6.08 sec (s.e. = 0.49sec). We carried out similar statistical tests for study 1A.

Figure 16. Task completion time in sec (left) and false targets selected (right) over number of states (horizontal axis) and visualizations (B / L = Bull’s-eye / Linear Visualization) for Study 1B

We found a significant effect of number of states on trial completion time (t = -3.422, p < 0.01) and number of false targets selected (t = -2.55, p < 0.02). Users were significantly faster and selected fewer false targets when performing with three states (mean 4.94 sec, 0.0 errors) than with five (mean 7.27 sec, 0.69 errors) as can be seen in Figure 16.
As with Study 1A, we found no significant effect of visualization on trial completion time ($t = -0.560, p > 0.5$) nor significant effect of visualization on number of false targets selected ($t = -0.221, p > 0.5$).

We found a significant difference between max(EI) and min(EI) measured in the eye-closed task and 3CVT over 17 participants of study 1A and 1B ($p < 0.001$). The average values of max(EI) = 0.81 (s.e. = 0.13), and min(EI) = 0.32 (s.e. = 0.02) (see Figure 17).

![Figure 17. Average max(EI) and min(EI) measured in the first two tasks over all participants of study 1A and 1B](image)

**3.5.4. Discussion: Study 1A and 1B**

Results from 1A and 1B reveal that the number of states is an important factor in the control of an on-screen cursor using engagement. Our studies showed that five states are preferable over ten states, and three states is significantly faster than five. However, three states offer less discretization flexibility: distances between levels must be greater than 0.33 or the start and target levels may sometimes overlap. We therefore chose five states as a comfortable range in which users can move a cursor. We use this value for our remaining studies.

Additionally, in our studies we noticed that the max and min values of the Engagement Index were similar for most participants. This might suggest that these values can be generalized to all users and can be used without calibration.

**3.6. Experiment 2: Engagement-based distance control**

Experiment one showed that users were successfully able to move the cursor in order to select items on a one-dimensional visualization. This experiment sought to investigate the effect of engagement control direction and distance in greater detail.
3.6.1. Experimental Design and Procedure

The interface, design, and procedure for experiment two are identical to experiment one, except as noted below. In this study we manipulated the selection distance factor.

- **Number of states**: From the previous two studies we saw that five states were optimal for this type of cursor movement and selection task. For this experiment, the number of states remained constant at five.

- **Visualization**: The first experiment showed that there was no significant difference between the selection time for the two visualizations, with study 1B finding the linear visualization was more accurate for selection. For this reason, we chose only the linear visualization.

- **Block design**: We used four blocks with each block sampling each of the eight distances two times.

- **Distances**: We selected four forward (increasing engagement) and four backward (decreasing engagement) distances: +0.2, +0.4, +0.6, +0.8 and -0.2, -0.4, -0.6, -0.8. Recall that in a five state condition a movement of 0.2 equates to shifting one state. Where possible, the start position of these distances was also varied.

3.6.2. Participants

Eight participants (5 female and 3 male) between the ages of 21 and 32 were recruited from a local university to participate in the experiment. All participants had heard of an EEG headset but none had used one before. They performed the eyes-closed task and 3CVT to obtain their minimum and maximum values of EI before performing the cursor control task.

3.6.3. Results and Discussion

The total number of trials that were successfully completed was 503 out of 512. The average trial completion time for these trials was 8.16 sec (s.e. = 0.41 sec).

As one would expect, an ANOVA showed that there was a significant effect of distance over trial completion time ($F_{7,49} = 5.014, p < 0.001$). Tukey post-hoc pair-wise comparison revealed the following pairs of distances were significant (-0.4, 0.6), (-0.4, 0.8), (-0.2, 0.6), (-0.2, 0.8), (0.2, 0.6), (0.2, 0.8).

Our results show that users are equally proficient in both directions (less to more engaged or more to less engaged state) as post-hoc comparisons did not reveal significant difference between pairs (-0.8, 0.8), (-0.6, 0.6), (-0.4, 0.4) and (-0.2, 0.2).
Figure 18 shows the mean task completion time for each distance. The graph shows an almost linear correlation between the distance moved and the time taken—$R^2$ values fall between 0.78 and 0.93. There are two important observations from this plot. First, the similar trend slopes indicate that the effort required for movement is similar in both directions. Second, while this was not a reciprocal tapping task, there is evidence from the plot to support the idea that Fitts’ law may be a good predictor of time performance with brain-controlled movement using EEG.

![Figure 18. Mean task completion time for experiment 2; trend line shows linear correlation between time (vertical axis, in seconds) & distance (horizontal axis)](image)

While in Study 1 participants found it easier to decrease engagement rather than to increase it, we did not find a similar affect in Study 2. In Experiments 1A, 1B and Experiment 2 the mean trial completion times for five states are similar for distances of -0.6 and 0.4. Note that Figure 16 (left) includes the mean time for both 5 and 10 states so it cannot be directly compared to Figure 18 but it gives an indication that the data is similar. We thus believe that any difference in performance in either direction is small and less important at five states than when a higher number of states are in use.

Finally, we found no significant effect of distance on the number of false targets selected ($F_{7,56}=1.029$, $p>0.05$). Figure 19 shows the mean false targets selected by distance.

![Figure 19. Mean false targets selected (vertical axis) for each distance (horizontal axis) with standard error bars for experiment two](image)
3.7. Experiment 3: Distractions & multi-tasking in a game setting

With a better understanding of the low-level performance of BCI input from experiments 1 and 2, we applied our guidelines to a gaming task. The goal of this experiment was to examine users’ performance in a situation that required multi-tasking.

3.7.1. Design

We designed a simple game that required players to complete two tasks simultaneously—one required the use of traditional inputs (keyboard and wireless remote), the other BCI-measured engagement. The user is required to move an on-screen avatar from the start point on the left, to the end point on the right. During this task, a demon sends a fireball towards the avatar, which the player must destroy by modifying their engagement to the required level. The trial is success only if the player destroys the fireball and reaches the end point within 30sec.

3.7.2. Experimental Interface and Task

The experimental game interface (Figure 20) was displayed on a 19-inch widescreen monitor (resolution 1280×1024 px). Participants sat 90cm from the screen.

![Figure 20. Visualization of the avatar (car), fireball (red circle), and engagement levels in case of not-matched (left) and matched (right)](image)

Each trial begins when the user reaches either the maximum or minimum engagement level (on a scale from level 0 to level 4), as visually indicated on the screen. This ensures all participants begin with the same engagement level. The participant presses a key on a keyboard to move the character toward the finish line. A demon appears and fires a fireball toward the avatar as it crossed the $\frac{1}{4}$ distance position. A linear visualization with five states of engagement is attached to the character. Each state has a number inside and it is coloured in a different shade of orange. The fireball has a shade of red and a number inside that matches one of the states on the visualization (Figure 11, left). When the user reaches the desired engagement level, its colour on the linear visualization changes to red to indicate the matching of engagement state with the fireball’s ‘kill’ state (right).
To destroy the fireball the user adjusts their level of engagement to match the number that is on the fireball. Once at that state, the user presses a key on the remote to destroy the fireball.

3.7.3. Experimental Design and Procedure

The experiment used a $2 \times 8$ between-participants design, with factors:

**Expertise**: Expert or novice user. Expert users had previous experience with the system through participation in either experiment 1 or 2.

**Distance**: +0.2, +0.4, +0.6, +0.8, -0.2, -0.4, -0.6, -0.8

The order of presentation of the distances was counterbalanced between participants. Each experiment consisted of four blocks of 16 trials. As with the previous experiments each participant was provided a full explanation of the study and given ample practice before the measured trials began. Users were encouraged to complete both tasks simultaneously - moving from one side of the screen to the other took 20 seconds, leaving only 10 seconds of extra time if the participant wished to pause to complete the engagement portion of the task (participants were informed of this). 10 seconds is about the average time we had previously observed for completing this task. Participants had three minutes of rest after each block.

3.7.4. Participants

Ten participants took part in this experiment. We used five participants who had previous BCI experience from one of our prior studies (4 male and 1 female) between the ages of 20 and 29. Their max(EI) and min(EI) obtained from the previous experiment were re-used. However, these values could change adaptively during the experiment if participants reached an EI higher than current max(EI) or lower EI than current min(EI). Another five participants (4 male and 1 female) between the ages of 22 and 27 had heard of an EEG headset but had not used one before. These five people completed the eye-closed task and 3CVT, then performed the game control task. All were recruited from a local university.

3.7.5. Results & Discussion

With 10 participants, the system recorded a total of 640 trials. Of the 320 experienced user trials, 302 resulted in successful completion of the task. In the remaining 18 trials the users always failed to kill the demon before reaching the end. Novice users failed to complete the trial on 43 occasions, again always due to a failure to kill the demon.
The average successful trial completion time, for all users was 21.2 sec. This was expected as the task was time-bound, with participants encouraged to complete the task in the time allocated. Repeated measure ANOVA showed a significant effect of distance on trial completion time ($F_{7,63} = 2.9$, $p < 0.05$). Post-hoc Tukey tests (HSD = 0.51, $\alpha = 0.05$) showed differences between the distance pairs: (-0.4, 0.8), (0.4, 0.8), (0.6, 0.8). Repeated measure ANOVA did not show a significant effect of experience ($F_{1,8} = 2.434$, $p > 0.05$) on trial completion time.

Repeated measure ANOVA on number of false targets selected showed no significant effect of experience ($F_{1,8} = 5.26$, $p = 0.051$), a significant effect of distance ($F_{7,56} = 8.17$, $p < 0.001$) and an interaction between the two ($F_{7,56} = 3.52$, $p < 0.05$). Experienced users selected significantly fewer false targets than novices (experienced mean = 1.8, s.e. = 0.9; novice mean = 4.3, s.e. = 1.6).

![Number of false targets and mean stop duration for Experienced (E) and Novice (N) users](image)

**Figure 21. Number of false targets and mean stop duration for Experienced (E) and Novice (N) users**

We counted how frequently users stopped moving the avatar to kill the demon and measured stopping duration to see if there are any differences in how frequently and for how long experienced and novice users stopped to complete the engagement task. Univariate ANOVA did not show any significant effect of experience on either of these measures. There was also no significant difference for stop duration. Figure 21 summarizes these results.

### 3.8. Discussion, Implications, and Applications in Game Design

BCI-based input has the potential to revolutionize gaming. The results presented in this chapter can guide design choices to enrich playing experiences.

#### 3.8.1. Continuous vs. Discrete Input Control

Our studies have shown that users can comfortably select from a list of five items (with a proportional increase in time as the selection distance increases). By modifying the discretization of the engagement range game developers can modify the difficulty of parameter
manipulation - increasing the number of levels requires more accuracy and greater practice and skill level. The studies also showed that one direction of selection is not favoured over the other and that a linear mapping of engagement to cursor movement is appropriate for this type of task.

3.8.2. Error Resilience in BCI Game Design

The swiftly changing nature of brain signals and the flickering inherent in BCI input due to this constant change makes very precise selections difficult. Unlike the common controller (keyboard, mouse) where a ‘slip’ error is a cause for incorrect input, BCI input is more erratic and influenced by environmental conditions. A player’s emotional involvement can easily spike when triggered by an external source (for example, a friend walking into the room).

Experiment two showed that longer distance selections are more difficult (take more time) than smaller, closer selections. In these situations, developers can implement cumulative functions to let users achieve goals in small steps, rather than giant leaps. For example, a user might be able to pick up a distant sword by achieving a very high level of engagement. If they cannot reach this high level, then the sword may edge closer and closer to the user as they sustain a lower level of engagement. This choice is captured in our discussion on relative input control presented above.

Game input needs to provide error tolerance to cope with this alternative and possibly noisy input stream. Triggering actions such as firing a gun or casting a spell should not be purely based on the user’s engagement level, as this could lead to unintentional actions. In this work, small-scale fluctuations were smoothed by the cursor becoming thicker and combined with users indicating via a button when their engagement level was appropriate to make a selection. This latter type of mode switching may not be appropriate for all types of gaming.

3.8.3. Lessons for Designers

We provide the following guidelines to designers: 1) For best performance discretize engagement into five or fewer discrete states; 2) Visual feedback showing the cursor, current state, and goal states should be clear to the user; 3) Engagement states tend to rapidly fluctuate and users should be allowed to achieve ‘hard to reach’ states in multiple steps; 4) Novice users require training, thus short games suited as ‘training wheels’ could prove beneficial to transfer learners into expert mode; 5) Combined input with a typical controller is possible for both walk-up and expert users.
3.9. Summary

The results of our experiments have led to a series of guidelines for designers of BCI-based games. We show that when discretizing raw engagement values five levels provide a good balance of speed and accuracy, users have good bi-directional control of engagement and are able to effectively control engagement in game-like multi-tasking scenarios. Our results also show that novice users are able to control engagement in a game like task and with a short training session they can noticeably improve their accuracy. Through detailed discussions and examples we show how our results could be used in game settings. The main contributions of this chapter are a systematic exploration of various parameters that influence mappings of engagement to virtual cursor control, a series of guidelines for designers and a demonstration of the practical value of our mappings in a gaming application.

As pointed out earlier, previously engagement was used to passively improve users’ performance during various tasks when their engagement in the tasks changed. However we want to investigate whether users can manipulate their engagement actively. If so, how do we quantify engagement and can it be an additional input for interactive applications are valuable questions for application designers. Moreover, practicing the ability to control engagement can help users engage in the tasks better, which improves productivity. Similar practice has been investigated with the Brainball game [88] where it was proposed as an interesting way of practicing control over brain activity and of learning how to relax.

We have presented scenarios and guidelines of how to use task engagement in an EEG-based interactive application as active BCI system. Users can actively control their task engagement successfully to communicate with the system, as the main input method and additional input method. In the next chapter, we will continue exploring task engagement, but in a passive BCI context. Specifically, task engagement can be monitored during task interactions in order to assist user as well as evaluate or compare interactive tasks.
Chapter 4. Measuring task engagement to evaluate interactive tasks

Figure 22. Examples of photo arrangements by Geotags (a); find someone’s photos (b); and photo corrected by meta-data (c)

We demonstrated in the previous chapter that task engagement could be used as an input for gaming applications. We also presented several guidelines so that designers can benefit by using our results. They were to demonstrate how to employ task engagement, calculated from EEG signals, to design an active BCI-based system. In this chapter, we move forward to investigate how to harness task engagement for a passive BCI-based system, where brain activity arises without the purpose of voluntary control. Consequently, in this chapter we employ the continuous usage of EEG to calculate task engagement of users while interacting with an application. Therefore, we first propose D-FLIP as an example of interactive application. It is a novel algorithm that dynamically displays a set of digital photos using different principles for organizing them. A variety of requirements for photo arrangements can be flexibly replaced or added through the interaction and the results are continuously and dynamically displayed. As a consequence, the global layout of all photos is automatically varied. We first present examples of photograph behaviours that demonstrate the algorithm and then investigate users’ task engagement using EEG in the context of story preparation and telling. This is organized to show how EEG can be used to evaluate an interactive application, with D-FLIP as an example, as a demonstration of a passive BCI system. The design of D-FLIP as a demonstration of an interactive application in HCI, however, does not count toward the contributions of this thesis.

Work described in this chapter was published in ACE ’13 as a full paper (Honourable mention) and in SIGGRAPH Asia E-Tech ’13 as a demo with title “D-FLIP: Dynamic & Flexible Interactive Display”.

52
4.1. The design of D-FLIP

Pervasiveness of digital cameras has led to large collections of digital photos that users often browse on computer displays, by rearranging them to gather similar ones based on specific features/ meta-data. Although several techniques exist to do this efficiently, most of these are somewhat systematic or goal-driven in terms of applying principles for displaying photos. These methods are useful in systematically organizing and finding photos but previous studies suggest that users often browse their photo collections without a specific search goal (e.g. [103]). Moreover, users often browse photos with actions such as enlarging displayed thumbnails in a certain order, displaying photos randomly on a digital photo frame or starting a slideshow for personal gratification and pleasure. To support these behaviours, the presentation of photos should flexibly and dynamically adapt to visual effects based on user’s input.

Additionally, one of the most enjoyable parts of personal photos is to share memories and reminisce with friends or relatives. Previous attempts to provide such experiences have revolved around presenting a static collection with interaction capabilities on tables [165] and handheld devices to facilitate story telling [21]. However, we want to improve not only the display layout but also the dynamic behaviours of the photos during interactions.

Therefore, we propose a novel method to flexibly display a set of photos by showing each of them in a dynamic and continuous motion like a living object. It allows users to replace or add displaying principles interactively and flexibly. As in Figure 22, users can manipulate (for example, enlarge and translate) a particular photo through flexibly grouping and arranging them using meta-data and/or their feature extracted values. In order to achieve such flexibility, we introduce an approach based on emergent computation. Geometric parameters (i.e. location, size, and photo angle) are considered to be functions of time. Photos are dynamically moved toward the directions determined by local relationships with adjacent photos at each time instance. As a result, the global layout of all photos varies automatically; converging gradually with time.

We illustrate example behaviours of photos and then do a user study to evaluate D-FLIP against Windows Explorer, a photo managing program familiar to Windows users. The evaluation involved two participants, a narrator and a listener to prepare and share a story. We measured both participants EEG signals to quantitatively measure users’ mental effort/ task engagement. In addition, NASA-TLX forms were collected from the narrators and listeners after each task. Our results show that with the same task, D-FLIP requires less task engagement and mental effort for story preparation.
4.2. Algorithm Overview of Dynamic Display of Photos

Each photo has three parameters: its position, size, and rotational angle. These parameters are considered as functions of time and are controlled to arrange multiple photos simultaneously on a display. The photo movement is shown by gradually changing the values of these parameters at every time instant.

The algorithm is explained by Eq. (1):

\[
dx/dt = f(\vec{x}) + \eta
\]  
(1)

Here, \( \vec{x} \) is a set of the three parameters above and its variation \( dx/dt \) is derived by \( f(\vec{x}) \), the principle to achieve the photo arrangement, and noise term \( \eta \). Larger amplitude noise increases the fluctuation and is useful for escaping local optima.

Furthermore, Eq (1) can be re-written in another form with the weight coefficients:

\[
\frac{dx}{dt} = \sum_i \{w_if_i(\vec{x}) + \eta_i\}
\]  
(2)

In here, \( f(\vec{x}) \), a variety of principles, is used to achieve the photos arrangement or layout. Let \( P \) represent the data of a photo, \( I \) represents the information of certain input or output devices, \( \vec{P} \) is all the photos in the environment, \( \text{Position}(P) \) is the photo position, \( \text{Size}(P) \) is its size, and \( \text{Rotation}(P) \) is its rotational angle. Assume that the number of principles related to position, size, and rotational angle are \( l \), \( m \), and \( n \), respectively. Eq. (3) is obtained by modifying Eq. (2). It controls the parameters of photo \( P \) and is calculated from all photos. Here, \( f_{Pl}(\vec{x}) \), \( f_{Si}(\vec{x}) \), and \( f_{Rl}(\vec{x}) \) are functions that represent the changes of position, size, and rotation, respectively:

\[
\begin{align*}
\frac{d}{dt}\text{Position}(P) &= \sum_i^n \{ f_{Pl}(I, \vec{P}) + \eta_i \} \\
\frac{d}{dt}\text{Scale}(P) &= \sum_i^m \{ f_{Si}(I, \vec{P}) + \eta_i \} \\
\frac{d}{dt}\text{Rotation}(P) &= \sum_i^n \{ f_{Rl}(I, \vec{P}) + \eta_i \}
\end{align*}
\]  
(3)

4.2.1. Principles of Photograph Arrangement

There are two types of principles that are important for photo arrangement: packing and mapping. Packing is a geometric problem concerning about arranging multiple photos with different sizes and rotational angles in a pre-determined area; it avoids overlaps and empty regions as much as possible. On the other hand, mapping is a semantic problem concerned with locating each photo based on its content and interaction with users. One example of packing is
the function to enlarge each photo as much as possible, but avoid overlaps with adjacent photos by translation, rotation, or shrinking. Another example is the function that moves photos toward the inside of the displaying window to avoid exceeding its boundary. Examples of mapping are functions that attract photos with the same tag, or to enlarge an interesting photo by a viewer, and so on. Here, each function can be established independently based on an individual principle as well as be implemented without paying attention to the global coordination. Certain feature values of each photo are assumed to be calculated and stored in the tag beforehand (e.g. to specify a person, taken location, etc.). By replacing or adding functions that correspond to the displaying principles, different photo arrangements can be achieved flexibly.

a. Geometric Packing

Here we explain principles related to geometric packing. First, the principle to avoid overlaps with adjacent photos is represented by Eq. (4). Here, $N$ is the number of photos; $\text{Avoid}(P, P_i)$ is $P$’s vector for escaping when $P$ and $P_i$ overlap. $\text{Adjacency}(P)$ is the set of photos overlapping with $P$.

$$f_{\text{translation}}(I, \vec{P}) = \sum_{l=1}^{N} \text{Avoid}(P, P_l) \quad \text{if } P_l \in \text{Adjacency}(P)$$

Second, a photo is moved toward the inside of the window based on Eq. (5) if its position exceeds the displaying window’s border. Here, $L, B, R,$ and $T$ are the left, bottom, right, and top coordinates of the window, $L(P)$, $B(P)$, $R(P)$, and $T(P)$ are the corresponding photo coordinates, and $A_l, A_b, A_r,$ and $A_t$ are their coefficients, respectively:

$$f_{\text{fold}}(I, \vec{P}) = \sum_{l=1}^{N} \begin{cases} A_l(L - L(P_l)) & \text{if } L(P_l) < L \\ A_b(B - B(P_l)) & \text{if } B(P_l) < B \\ A_r(R - R(P_l)) & \text{if } R(P_l) > R \\ A_t(T - T(P_l)) & \text{if } T(P_l) > T \end{cases}$$

Figure 23 illustrates how photos avoid overlapping. Without overlaps, each photo becomes larger until it reaches the predetermined maximum scale when Eq. (6) is applied (Figure 23a). If two adjacent photos overlap, the larger photo becomes smaller until it reaches the predetermined minimum scale when Eq. (7) is applied (Figure 23b); they move to opposite directions when Eq. (5) is applied (Figure 23a), or rotate in opposite directions when Eq. (8) is applied (Figure 23c). Here, $A_{l1}$ and $A_{l2}$ are coefficients, and $\text{Ang}(P_l, P_j)$ is the rotational angle by which $P_l$ and $P_j$ avoid overlapping:
\[ f_{\text{enlarge}}(I, \vec{P}) = \sum_i A_{sz}(\text{Scale}_{\text{max}} - \text{Scale}(P_i)) \quad \text{if} \ \text{Adjacency}(P_i) = \emptyset \wedge \text{Scale}_{\text{max}} \succ \text{Scale}(P_i) \]  

\[ f_{\text{shrink}}(I, \vec{P}) = \sum_i A_{s1}(\text{Scale}_{\text{min}} - \text{Scale}(P_i)) \]

for all \( P_j \in \text{Adjacency}(P_i), \text{Scale}(P_j) < \text{Scale}(P_i) \wedge \text{Scale}_{\text{min}} < \text{Scale}(P_i) \)  

\[ f_{\text{rotation}}(I, \vec{P}) = \sum_i \text{Ang}(P, P_i) \]  

The upper-right photo in Figure 23b will become as large as possible by referring to environmental parameters indicating the positions and sizes of adjacent photos. However, when two photos collide, the larger one becomes smaller, as shown in the lower-right photo based on Eq. (7) if these two equations are simultaneously applied. Thus, all photos are gradually arranged without empty space, and at the same time, their sizes become almost equal if there is not a change of conditions. Even if these two principles conflict, the algorithm will find a solution. Other principles related to geometric packing can be obtained similarly.

![Conceptual behaviours of photographs](image)

**Figure 23.** Conceptual behaviours of photographs

**b. Semantic Mapping**

As one of the simplest examples of semantic mapping, a function to enlarge an interesting photo is represented by Eq. (8). Here, Attention is a set of interesting photos given by adequate input device such as a mouse or a gaze input device:

\[ f_{\text{attention}}(I, \vec{P}) = A_{a1} \cdot \{\text{Scale}_{\text{max}} - \text{Scale}(P)\} \quad \text{if} \ P \in \text{Attention} \]

Eq. (10) shows how a focused photo attracts other ones with similar attributes. Here, \( \text{Similarity}(P_i, P_j) \) is the similarity between photos \( P_i \) and \( P_j \), and if this value is larger than a threshold, \( P_j \) moves toward \( P_i \), and away otherwise. The similarities are assumed to be calculated by feature values obtained by image processing or from tags of photos. Other related principles of semantic mapping can be obtained similarly.
\[ f_{\text{attraction}}(t, \vec{P}) = \\
\left\{ \begin{align*}
\sum_{i} A_{22}(\text{Similarity}(P, P_i) - \text{Threshold}) \cdot (\text{Position}(P_i) - \text{Position}(P)) & \text{if Similarity}(P, P_i) \geq \text{Threshold} \\
\sum_{i} A_{22}(\text{Similarity}(P, P_i) - \text{Threshold})/(\text{Position}(P_i) - \text{Position}(P)) & \text{if Similarity}(P, P_i) < \text{Threshold}
\end{align*} \right. \]  
(10)

c. Viewer’s Interactions

Even after the system reaches the balanced condition, the photograph behaviours can be observed when parameters of the display environment vary (e.g., when new photos are added or the size of the displaying window is changed). Also if a cursor (operated by mouse, gaze input, or other devices) is used, the photo overlaid by a cursor becomes larger using Eq. (8) with certain weight coefficients.

Users can observe the displayed photos on a computer display and interact with them at the same time by using cursors, touch panels, and gaze input devices. Multiple users may interact simultaneously with separate input devices. Keyword inputs, which are detected by a speech recognition technique, may also be an input method. On the other hand, the display resolution, the number of displays and their positions and orientations, and sizes are variables of the output. Such information about input and output is inferred in previous equations.

4.2.2. Behaviours of Photographs and Performance

a. Behaviours of Photographs

The experimental system was developed in C#, as an integrated development environment, and Windows User API, Microsoft .NET Framework 3.5, DirectX SDK (November 2008), and Microsoft XNA Game Studio 3.1.

Some photograph behaviours using the principles explained in the previous section are illustrated here and show the flexibility of our method. They are classified by types: geometric packing and semantic mapping.

b. Photograph Behaviours with Geometric Packing

Figure 24 shows how photos avoid overlaps when four equations (Eqs (4)(5)(6)(7)) are applied, and there are no interactions or other changes to environmental parameters. Figure 24 (left) shows the initial state where 75 photos are located randomly and overlapped. However, they gradually move to avoid overlaps and occupy the empty regions (using Eqs. (4)(5)), as shown from left to right in Figure 24. At the same time, the photos’ sizes vary (using Eqs. (6)(7)), then soon become almost equal (Figure 24 right).
Figure 25 is an example of photos avoiding overlaps by rotating in addition to the equations used in Figure 24. In Figure 25 on the left is the original layout, and on the right is the layout with collision-free arrangement with rotation. This is useful when photos are shown on tabletop surface displays shared by several users.

![Figure 24. Geometric packing of photos without the rotation](image1)

![Figure 25. Geometric packing of photos with the rotation.](image2)

Figure 26 shows photograph behaviours when one of the environmental parameters, the window size, is changed. Once the window is enlarged (left), the contained photos steadily move to the empty space (middle), according to Eq. (5). This is gradually convergent with time so that the sizes of all photos become almost equal but avoid overlapping (right).

![Figure 26. Sequence when size of the displaying window is changed](image3)
c. Photograph Behaviours with Semantic Mapping

Figure 27a shows an example where one photo is focused by overlaying a cursor (at the bottom center of the photo). Figure 27b shows photos arranged by colour and user interests using the principles of geometric packing (i.e., Eqs. (4)(5)(6)(7)). In this figure, there are two cursors (magenta and green) pointing at two photos (bottom-left night scene and upper-right daylight scene, respectively). Soon photos with similar colours are moved toward the focused ones. The final layout is achieved by Eq. (9) with semantic mapping principles. Similarly, other feature values calculated by image processing can be used to a group of photos by applying this principle. Figure 27a shows an example of photos arranged using Geotags. In this example, photos arranged without overlapping (using the Eqs. (4)(5)(6)(7)) are attracted by their geographical identification metadata (latitude and longitude coordinates) based on Eq. (9) and moved to their corresponding positions on a world map.

![Figure 27. Examples of geometric packing where a photo is focused by overlaying a pointer (a); and semantic mapping where photos are arranged by colour and interest similar colour photos are collected (b)](image)

Figure 22b shows an example of finding someone's photos. When a human face in a photo is selected, the size of that photo becomes larger by Eq. (8). Also, all photos containing the selected person are attracted and gathered around the focused photo dynamically by Eq. (9) given a photoset of a social relationship. Here, a face recognition is assumed to be working and tags for the faces are adequately given in advance. Similarly, Figure 22c displays examples of grouping photos with closed curves drawn by a mouse using meta-data given to each of the photos in advance. In this figure, photos having meta-data of Mr. A and Mr. B are gathered in the red closed curve in the left and the green closed curve in the right, respectively. In the overlapping area of these two closed curves there are photos belonging to both Mr. A and Mr. B.

4.3. Discussions

A variety of photo arrangements can be achieved by replacing or adding functions without explicitly defining any prior relationships. As a result, even though the principles may conflict,
our algorithm can always find a solution. In addition, compared with approaches based on combinatorial optimization, which always try to find a global optimal solution, our method generates solutions at every instant without fearing the local optimal solution. This is because we do not have to try finding the global optimal solution in the environment where conditions always dynamically vary.

Theoretically, the displayed photos by our algorithm are slightly vibrated because the function inherently includes a noise term in Eq. (1) that causes the system to be constantly in motion. In the examples described above, a fixed small amplitude was used for the noise. Larger amplitude noise increases the fluctuation of the environmental parameters, and such fluctuation is often useful for escaping local optimum. But if such large vibration is not suitable, the vibrations can be removed using a filter before rendering. In addition, the amplitude can be varied during photo arrangement to obtain a better performance.

4.4. Performance Evaluation

In order to know the computational performance of the proposed algorithm, the frame rate of the implemented system was evaluated with various numbers of photos, and the influence of principles on the performance was analysed. Here, the experimental system was run on a Windows 7 64-bit PC with Intel Core i7 CPU (3.20 GHZ) and 12.0 GB memory. It was not implemented for parallel processing; therefore, it ran mainly on a single core of CPU. The graphics card used was NVIDIA GeForce GTX 285.

![Transition of frame rate](image)

**Figure 28. Transition of frame rate**

Regardless of the initial or convergent state, the computational complexity of the proposed algorithm is expected to depend on the number of photos and the principles of photo arrangement. Thus, we computed the frame rates under various conditions (various numbers of photos and principles applied) and analysed the factors that influence the performance of our
method. In this experiment, standardized JPEG photos with 512x320 pixels were used, and their display area on the screen was 1024x768 pixels. The following four kinds of equations were denominationally applied: Eq. (4) for translating each photo to the opposite direction when a collision occurs, Eq. (5) for moulding each photo not to exceed the window's border, Eq. (6) for enlarging each photo as much as possible without overlapping, and Eq. (7) for shrinking the larger one to avoid overlapping.

Average computation times (FPS, frame per second) for each ten-second processing cycles are plotted in Figure 28. In this graph, the cyan line represents the result when all four principles are switched on. Magenta, blue, and green lines represent results when one of the four principles is switched off; the magenta line represents result with Eqs. (4)(5)(6), the blue line represents result with Eqs. (5)(6)(7), the green line represents result with Eqs. (4)(5)(7), and the red line represents result with only Eq. (5), respectively.

Note that the performance was saturated at around 60 FPS for a smaller number of photos (less than 600 or 700). However, when the number of photos exceeds 700, the performance drops rapidly. From these results, notably better performance can be obtained if Eq. (4) for translation and Eq. (6) for enlarging are not used because probabilities of photo overlap decrease. At the same time, the performance was deteriorated without using Eq. (7) for shrinking because the overlapping occurs more frequently. Therefore, we can conclude that the overlapping avoidance is crucial for the performance of our algorithm. In addition, even when only the equation for moulding (Eq. (5)) is active, the frame rate is not so high when the number of photos becomes larger (see red line in the graph), because the system needs to determine intersections between each of the photos and border of the displaying window.

4.5. Measure and process EEG signals to calculate task engagement

4.5.1. The Effect of Animation on Users’ Interest

Compared to static method, dynamic photo displaying seems to be more interesting and aesthetically appealing [99]. Previous studies have shown that animation can boost users’ performance in learning and teaching for example when understanding Newton’s law of motion [157]. In other words, animations can help users perform the task (e.g. learning) more easily. In terms of users’ interest, both static and animated graphics can increase interest. Animation is likely to increase emotional interest (created by events that are arousing) while static graphics are likely to trigger more cognitive interest (related to the connections between incoming information and background understanding [102]). As the result, D-FLIP is likely to trigger
emotional interest from users because of its dynamical and interactive movements. This will help to achieve the goal of D-FLIP, which is letting users view photos interactively with ease and interest.

4.5.2. Evaluating using Neural Signals

Traditionally, interactive programs are evaluated by investigating performance (i.e. [160]), or users’ behaviours. However, with programs designed for easy use and pleasure, an evaluation method measuring users’ affective or inner states is more suitable. Usually, this is done by questionnaires answered by participants during the experiment. However, this takes place after the event when important issues may be forgotten. Additionally, participants may not be aware of their states or might simply guess. Neural signals, measured from the brain can better reflect a users’ current state and provide an evaluation metric.

There are many methods to detect neural signals such as fMRI, MEG, functional near-infrared spectroscopy (fNIRS) and Electroencephalography (EEG). A brief summary of those techniques is discussed in [184]. In addition, EEG devices are portable (compared to fMRI, MEG) and have high temporal resolution (compared to fNIRS). EEG signals have been also shown to capture the affective state (such as relaxation [88], arousal [85], and task engagement [27]).

One purpose of our proposed algorithm, D-FLIP, is to help users browse photos with ease and interest thus measuring task engagement. According to Matthews et al. [125], task engagement can be defined as the effortful concentration and striving towards task goals where task demands and personal characteristics may influence this pattern of processing. In another work, Matthews et al. noted that effort investment in the task would increase task engagement [126]. Studies have shown a positive correlation between EEG engagement and task demands including stimulus complexity processing and the requirement for attentional resources allocation [25, 171]. EEG engagement correlated with task loads in not only simple vigilance and memory tasks but also in complex simulation tasks such as in radar operations simulation environment [25, 26]. Consequently, if an application requires low level of task engagement in using, it also requires low levels of task demands, task loads, and workloads. In other words, it can be considered easier to use when compared to applications requiring higher task engagement.

Pope et al. [149] present a measurement of task engagement from EEG as $\beta/(\alpha + \theta)$. Further studies [71, 150], with the same measure were used to show the performance quality benefits and reduced mental workload. These positive findings have been also replicated using
extended periods of task performance [70] and a vigilance task [129]. Given this evidence of measuring users’ task engagement / workload, we used that formula to evaluate our system by comparing D-FLIP with a similar and competitive program.

4.5.3. User evaluation of d-flip

Different photo arrangements with dynamic motions, shown in previous section, are expected to be effective in many situations such as viewing many photos at once, surveying a set of photos from different viewpoints (layouts), and finding pictures from the dynamic motions. Also, the smooth visual effect caused by interactions will keep up users’ motivation to actively look through and interact with photos. We designed a user study to investigate this further.

We compared D-FLIP to the Windows Explorer, a default Windows program as it can be easily customized to include a storyboard area in order to support storytelling. Further, Windows Explorer is familiar and supports basic functions with photos such as viewing a large collection at once, enlarging, and arranging photos (along with PhotoMesa [23], Picasa (http://picasa.google.com/), etc).

a. Task Design

The experiment is a modified version of [158] where a user shares a story with a friend by showing her the collections.

Each experiment session required two participants: one narrator and one listener. The narrator brought two sets of photos that were taken by them from different events. The participants sat beside each other in front of a 27-inch monitor (2560x1440) displaying the narrator’s personal photos. The narrator then examined the photos and selected 10, which he thinks, represent his/her story. Afterward, he narrates the story to the listener. The narrator was encouraged to use the selected photos to tell the story but could also use other non-selected photos to enrich it. The narrator was encouraged to tell about the actual events that happened in the photos as well as to interact with the program in a natural manner. During these steps, the listeners observed, listened, and enjoyed the story. We measured the task engagement from both narrator and listener to investigate the effect of the interactive application on the person who actually interacts with the program and on the person who does not interact but observe the interaction and listen to the story. Each participant wore an Emotiv EEG headset during the study.

There is a storyboard in D-FLIP to support the storytelling mode. The area has ten equal boxes, which hold ten selected photos (Figure 29, left). In case of Windows Explorer, we used two
Windows Explorer windows with one window above the other. The bottom window was used to store 200 photos and the top window was used to store the selected 10 photos (Figure 29, right). In Windows 7, the layout of these windows was modified to display all photos in thumbnails and without filenames. All bars and additional panes (e.g. Library, Details, and Navigation) were hidden to make Windows Explorer comparable to D-FLIP in terms of visualization and functions. Features of D-FLIP in this experiment included dynamic arrangement when overlaying a pointer; attraction for photos with similar colours; and timeline to display photos in chronological order (enabled manually by the narrator). These features were chosen as they are comparable with features of Windows Explorer. Moreover, the feature which enables similar colour photos gathering was chosen because of the assumption that user usually take several photos of one scene in order to choose the best one afterward. Hence there are many collections of similar photos in each photoset.

b. Method

14 participants (5 females and 9 males) between the ages of 19 and 30 participated in the study and were arranged into 7 pairs. All were from the local university and had at least 5 years experience of using Windows Explorer for viewing photo collections. The narrators brought to the experiment their 400 photos divided into 2 sets of 200 each. These photos were resized to 640x480, rotated if necessary by the experimenter before beginning the study. Most narrators brought photos from their previous trips (e.g., a game exhibition or scenic photos of various places). All participants had normal or corrected-to-normal vision. The entire experiment took about 1 hour 15 minutes and participants were paid for their time (about 10 USD).

![Figure 29. Storytelling mode with D-FLIP (left) and Windows Explorer (right)](image)

First, the narrator had adequate time to practice with D-FLIP and Windows Explorer using a sample photoset (Figure 29). After this they wore the Emotiv headsets. Each experiment session had two blocks with either D-FLIP or Windows Explorer. Each block started with a 3-minute reference recording where the participants sat still, comfortably, and with eyes open. After this,
the narrator prepared a story in 5 minutes by selecting 10 photos from her first photoset. The narrator then told the story to the listener within 5 minutes. Both participants had a 2-minute break before the same procedure was repeated with the next program. NASA-TLX forms were given to both participants after each block. These forms collected ratings of mental demand, physical demand, performance, effort, and frustration. The order of the two programs was balanced between experiment sessions.

In the storytelling step, we only recorded EEG signals from the listener. This was because the narrator needed to speak freely with facial, hand, and body movements which would contaminate the EEG signals. To reduce the noise in EEG data (we also decontaminate the signal) we asked: (1) the narrator to not wait for feedback from the listener; (2) the listener to minimize body or facial movements as feedbacks to the story; and (3) both participants to focus on the monitor. The narrator was encouraged to use both programs in a natural manner.

c. Data Acquisition and Analysis

We used an Emotiv EEG wireless headset for measuring EEG signals. Its signals quality has been validated by previous studies (e.g. detecting and classifying Event Related Potentials [38, 184], evaluating visualization effectiveness [14], exploring the nature of decision making [80], and designing a BCI controlled video game [9]). This headset is portable, easy to setup, and most users can wear it comfortably for at least an hour [58].

EEG signals were sampled at 2 kHz internally then down-sampled to 128 Hz output. We inserted three types of markers into the marker channel which is in parallel with EEG signal channels. These were: start recording, stop recording, and eye blinking events. We then adapted the “sliding window technique” which is commonly used to process EEG signals (e.g. [13, 113]). The recorded EEG signals were segmented into 4-second epochs with 2-second overlaps between them. The DC offset was removed by subtracting the averaged value from the reference signals.

d. EEG Signals Decontamination

To remove artifacts from muscle movements and eye blinks, we adapted a decontamination process [25] to decontaminate Emotiv’s EEG signals. These steps are: (1) Three data point spikes detection with amplitudes greater than 40uV (e.g. caused by tapping or bumping the sensors); (2) Notch, low and high pass filters to remove unwanted artifacts (e.g. from muscle movements); (3) Detect and remove eyes blinks. To aide in step (3), we used Emotiv SDK to
insert markers when a user blinked. Two independent marker channels for the narrator and listener were used to remove eye-blinks [25].

We then performed a validation step where epochs were removed if the combined power of the alpha, beta, or theta frequency bands of an epoch changed more than 20% of its original values. We removed totally 7.83% of collected signals out of which 5.13% due to excessive signals and 2.7% due to high changes in combined power spectral.

e. Task Engagement Calculation

Engagement index was calculated as in Pope et al. [149]. $\beta$, $\alpha$, and $\theta$, is the combined power in the ranges of 13-22Hz, 8-12Hz and 5-7Hz frequency bands.

$$Engagement\ Index = \beta/(\alpha + \theta)$$  \hspace{1cm} (11)

We used F3, F4, FC5, FC6, P7, P8, O1, and O2 channels as they are the surrounding channels of Cz, Pz, P3, and P4 which were used by Pope et al. [149]. Engagement indexes were calculated in 4s window for each task & user separately and then normalized for the narrator or listener.

f. Results

Table 2 summarizes the results, which are the averaged task engagement during the session, for two types of tasks (Story Preparation and Story Telling), two types of programs (Windows Explorer and D-FLIP), and with two participant types (Narrator and Listener).

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Programs</th>
<th>Narrator</th>
<th>Listener</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story Preparation</td>
<td>Windows Explorer</td>
<td>0.603</td>
<td>0.480</td>
</tr>
<tr>
<td></td>
<td>D-FLIP</td>
<td>0.554</td>
<td>0.436</td>
</tr>
<tr>
<td>Story Telling</td>
<td>Windows Explorer</td>
<td>0.512</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D-FLIP</td>
<td></td>
<td>0.440</td>
</tr>
</tbody>
</table>

Table 2. Average task engagement for narrators and listeners.

Figure 30 (left) shows a time series sample of engagement during the preparation task of one narrator. Figure 30 (right) shows the task engagement values of narrators in each experiment session for each of the two programs. T-test shows that there is a significant difference between those two programs ($t = 2.95, p < 0.05$). However, we found no significant differences between the types of program in listeners ($p > 0.5$). The details are shown in Figure 31 (left - Story Preparation) and Figure 31 (right - Storytelling). This may be due to the individual differences.
where the listeners did not directly interact with the programs hence the measured task engagement mainly depended on the storytelling skills of the narrators.

**Figure 30.** An example of a narrator's task engagement changes in time in one session (left); and Task engagement of each narrator (right).

**Figure 31.** Task engagements of listeners in the Story Preparation task (left) and Story Telling task (right) between Windows Explorer and D-FLIP.

From the NASA-TLX, we found a significant difference in mental demand of the narrator between two tasks: Windows Explorer (mean: 9.14) and D-FLIP (mean: 5.00). However, we found no significant difference in physical demand, temporal demand, performance, effort, and frustration of narrator between two tasks ($p > 0.5$). For the listener, no significant differences were found in mental demand, temporal demand, effort, and frustration between two tasks.

The results showed that the narrators had higher engagement when using the Windows Explorer compared to D-FLIP. This also means that Windows Explorer requires higher task loads hence users require more effort to complete the same task in the same amount of time compared with D-FLIP. Consequently, besides the benefit of having more interest, D-FLIP can help users perform the task easier compared to a common program such as Windows Explorer. However, we only found the significant difference in the narrator’s task engagement, not the listener’s. Obtained results from the NASA-TLX also showed consistent results compared to the captured EEG signals. A possible reason is because only the narrators were the ones actually interact with the programs while the listeners were not. Consequently, with less task engagement or
workload required in performing the task, users will find out that interacting easier and more enjoyable.

There is one possibility that the narrator were distracted by the continuous moving of photos when interacting with D-FLIP hence the lower task engagement. However, it is also possible that D-FLIP is harder for narrator to find a specific photo or to perform manipulating tasks (e.g. select, zoom in/ out) due to continuous movements of photos. This would cause a higher task engagement in using D-FLIP than in Windows Explorer. Moreover, although adequate practice time was given to the narrators prior to performing the task, there was possibility that the users felt more familiar with Windows Explorer and performed the task more skilful than D-FLIP. As a result, the actual difference between measured task engagements of two programs might be larger if another program was used instead of Windows Explorer (i.e. a less famous program such as PhotoMesa). Our results show that even if there are effects of distraction and familiarity, D-FLIP still requires less task engagement hence less task loads and mental demands.

4.6. Discussion

The user evaluation points out that D-FLIP requires a lower level of mental effort and task engagement compared to Windows Explorer. Several factors of D-FLIP may contribute to this lower produced task engagement such as: high visibility with variety of layout, dynamic and smooth motions of photos, gathering photos based on similar attributes. As a result, D-FLIP requires less effort from users to perform the task hence they can focus on the story contents and interactions with other users.

It can be said that lower investment of effort and task engagement in an interactive task also means that the task is easier to perform. However, if these values are too low users may feel bored and lose interest. Equivalently if the values are too high it can cause anxiety [41]. Our results show that D-FLIP helps users perform the task easier than a common program such as Windows Explorer. However, while taking advantage of this simplicity, D-FLIP still keeps motivation and interest in users due to the dynamic, flexible, and interactive motions of photos produced by the proposed principles. This combined with the fact that animated graphics enhances emotional connection with the system, makes D-FLIP a particularly powerful system for visualizing large collections of images.

Through performance evaluation, our proposed algorithm can work well with at least 700 photos. However, this can be easily improved further with a better computing hardware and parallelization for optimized implementation.
The smooth motion of each photo is characterized by (1) the implemented photo arrangement principles and (2) parameters values used in the equations (Eqs. (1) - (9)). These are implemented by careful considerations and preliminary experiments. Our next step is to improve the algorithm by categorizing the parameters and optimizing it for different types of interactive applications. Doing this will let the algorithm easily and flexibly adapt to various types of content. Besides existing content types (i.e. text, broadcasts, movies, web content, music), our algorithm can adapt to work with new and emerging types of content such as dreams visualization [92] where values of parameters are from fMRI patterns for an image. The images depicting a person’s dream will have smooth and dynamic motions, which make our visualization well suited for visualizing this type of data. This can also help to build an ecosystem of photos, which includes various promising features such as pigeonholing, printing, automatic acquisition of meta-data, evolving into photo sharing sites, and coordinating with social network services. We can also explore other interactive devices such as digital tables with multi-touch interaction, voice recognition, brain-machine interface, range image sensors, and other sensing devices to further enhance the fluidity of interaction in different application contexts.

Although the focus of this chapter is not on a novel evaluation methodology, we believe that our way of measuring task engagement using EEG offers greater insights into workings of an application. Our measured task engagement is consistent with the NASA-TLX results; providing a source of external validity to our measurement mechanism. This evaluation method can be used in other settings for similar comparisons. It can also be improved to capture emotions (e.g. relaxation, meditation) and users’ inner states (e.g. error awareness).

4.7. Summary

At the start of this chapter, we aim to design an interactive application and propose an evaluation method using EEG. Consequently, we first propose the design of D-FLIP, which allows users to dynamically and interactively view and manage their personal photos. It is presented as a common interactive application where the user’s affective states are highly interesting to the task designer.

Task engagement was calculated from EEG in a passive and background setting. It was demonstrated that task engagement could be used as a method to evaluate interactive applications. This example of the continuous usage shows that we can take advantage from
portable and commodity EEG headsets. This provides a powerful evaluation method to design interactive tasks in an insightful way and keep them in the balanced level of task engagement.

In the last two chapters, we have explored the capabilities of employing EEG using commodity headsets (such as Emotiv) in a continuous measurement and calculating task engagement. It can be used as an input control or as an evaluation tool for interactive tasks. In the next chapters, we will continue exploring passive BCI systems where EEG signals are captured in the background to potentially enrich the interaction. Additionally, we will explore the usage of commodity EEG headsets through ERP, a brain pattern elicited in certain task conditions. Although the existence of ERPs has been investigated thoroughly, it is un-explored in how to harness their existence in interactive tasks. We will investigate this problem and aim to provide guidelines to task designers by investigating Error-Related Negativity, a brain pattern associated with human error awareness.
Chapter 5. Error-Related Negativity for Single User Interactive Applications

This chapter examines the ability to detect a characteristic brain potential called the Error-Related Negativity (ERN) using off-the-shelf headsets and explores its applicability to HCI. ERN is triggered when a user either makes a mistake or the application behaves differently from their expectation. We first show that ERN can be seen on signals captured by EEG headsets like Emotiv when doing a typical multiple choice reaction time (RT) task – Flanker task. We then present a single-trial online ERN algorithm that works by pre-computing the coefficient matrix of a logistic regression classifier using some data from a multiple choice reaction time task and uses it to classify incoming signals of that task on a single trial of data. We apply it to an interactive selection task that involved users selecting an object under time pressure. Furthermore the study was conducted in a typical office environment with ambient noise. Our results show that online single trial ERN detection is possible using off-the-shelf headsets during tasks that are typical of interactive applications. We then design a Superflick experiment with an integrated module mimicking an ERN detector to evaluate the accuracy of detecting ERN in the context of assisting users in interactive tasks. Based on these results we discuss and present several HCI scenarios for use of ERN.

5.1. Introduction

An ERN is a form of an Event Related Potential (ERP) that can be triggered in the brain when a user either makes a mistake or the application behaves differently from her expectation. This pattern is produced in a person’s brain when she is aware of the obvious error(s) that s/he has made; either through system feedback or individual realization [68]. For example, ERN would be produced when pressing the LEFT key while intending to press the RIGHT key in a multiple choice RT task. It also appears, but with lower amplitude when a person is confused about the decision that s/he has made [47]. Usually these ERN signals appear and peak within 150ms of the committed action [73].

If ERN signals can be detected during an interactive task they can be used in detecting and correcting errors or in augmenting users’ experiences in those activities. An ERN detection

---

3 Work described in this chapter was published in CHI ’12 (Best paper award) with title “Detecting Error-Related Negativity for Interaction Design”.

71
module provides another medium for HCI designers to access users' intentions, which is intuitive and directly from the users' brain. This has great potential in many types of interactive application such as in gaming, spatial navigation tasks and aiding object selection. For example, when ERN is detected the system can prompt the user to check if the selection is the intended target.

Figure 32. A user performing a Superflick [154] while wearing an off-the shelf EEG headset to see if ERN detection can aide in an object selection task

However, it is difficult to detect a clear ERN pattern due to noisy EEG signals and lack of effective real-time algorithms. Most research in detecting ERN is focused on being able to detect it over an average of multiple trials relying on offline methods (see [47, 130] for examples). As ERN is known to appear when using multiple-choice RT tasks, most research is done on Flanker task [61]. Flanker task is a visual experiment where the participant has to respond to a central and directed symbol that is surrounded by distracting symbols. There are very few ERN detection algorithms that work in real-time as well as on multiple-choice RT tasks that are not Flanker tasks.

Additionally, most ERN experiments are performed using expensive, non-portable sensing devices such as NeuroScan from NeuroScan™, g.BCIsys from g.tec™ or ActiveTwo from Biosemi™. These devices have the benefits of not only more sensing channels (up to 256 electrodes that can cover all channels in the 10-20 international system), but also higher and user selectable sampling frequencies of up to 16 kHz. These systems also introduce less noise in the signals because of better sensing electrodes and integrated amplifier/converter module.

In comparison off-the-shelf EEG headsets have at best 14 channels, a sampling frequency of about 1 kHz and introduce more noise due to wireless transfer of EEG signals. Due to these limitations it is not clear if these devices can capture ERN patterns. However, due to their price and portability their use is being increasingly explored within HCI. In this chapter we
investigate the appropriateness of these off-the-shelf EEG headsets in detecting ERNs and also the effectiveness of real-time algorithms in doing so on a single trial and in real-time.

We first describe an online single trial ERN detection technique that is verified using data acquired from the frontal-central cortex of the human brain. We then show that we can detect ERN online on a single trial in an object selection task. This demonstrates the abilities of harnessing ERN in interactive applications in office conditions. Both our experiments show that we can detect ERNs using the Emotiv headset with an accuracy of up to 70%.

These rates are indicative of the type of accuracy one can expect from off-the-shelf EEG sets. Improvements in learning techniques can improve this accuracy but it is unlikely that detection accuracy will reach 100%. It is therefore expected that with high detection accuracy users might become over-resilient on a system detecting ERNs, which can increase the cost of recovery from the error. In order to examine this issue we conducted a final study where we compare users' error-rates when performing Superflick [154], a modified pointing task, with different potential ERN success rates. The results of the experiment show that ERN detection rates of 65 to 80% are acceptable for interactive applications. We finally discuss the implications of our results to interactive applications.

The contributions of this chapter are: (a) we demonstrate that off-the-shelf EEG devices such as the Emotiv headset can measure ERN from channels in the front-central part of the brain; (b) we demonstrate that it is possible to detect these ERN signals online (as opposed to offline) from a single trial in a task that is closer to the types of tasks encountered in HCI; (c) through a final experiment we show that detection accuracies in the region of 65 to 80% are sufficient to use these techniques in real-time interactive applications.

5.2. Error Related Negativity (ERN)

ERN is a pattern observed when a user makes an error in a reaction time task. Its shape is a negative deflection which appears in the on-going EEG right after the time the decision was made. The ERN also appears when users have feedback about their response accuracy [89]. Its amplitude is large when the user is clearly aware of his/her error and is small when user is confused (where errors are caused by data limitation) [47]. However, despite the change in amplitude, ERN latency seems to be consistent (about 100ms after the event). Interestingly, the amplitude of ERN does not depend on the behaviour accuracy itself but the user's perception of it [47, 195].
To date most of experiments involving ERN detection are done offline by averaging over multiple trials. Participants are asked to perform a multiple-choice task in which trials with incorrect responses were used to archive a clear ERN pattern. For example, Gehring et al. [75] used this methods to determine the effect of a speed/ accuracy trade off on different representations of ERN. Scheffers and Coles [47] use this method to conclude that ERN is a manifestation of the on-going monitoring system in the brain which compares the expected response and the actual response.

However, in an interactive application, ERN will be most useful when it can be detected immediately after it happens. This requires the online detection of the ERN pattern (as opposed to offline) from EEG signals. One way to achieve this is by caching and averaging a small number of patterns in a limited time in order to refine a reasonably clear ERN [166]. This method requires a waiting time to collect sufficient signals and therefore delays the progress of error correction as it goes through multiple trials.

Various attempts have also been made to detect ERN from a single trial. One example is from Ferrez and Millan [68] where they train a Gaussian classifier to recognize error and correct trials on a single trial basis but after the experiment was completed (so considered an offline single-trial approach). Another example is from Dal Seno et al. [49] who used a Linear Discriminant Analysis (LDA) classifier to cross check with every spelled letter using a P300 speller. The reported performance of the online version varies from 58% to 69%. This result is just better than chance but still encouraging, as this is one of the few attempts at detecting ERN online and from a single trial.

All the above experiments have been carried out on EEG headsets with up to 256 electrodes and with sampling frequency up to 16 kHz. Most importantly these approaches rely on having access to Fz, Cz, Pz and Oz channels. We are not aware of any attempt to detect ERN from a low-cost portable headset where the number of channels and sampling frequency are limited. Hence, we investigate if a low cost, features limited EEG headset can detect ERN and benefit to HCI.

5.3. Online ERN Detection

The online, single trial ERN detection algorithm we present below is an adapted version of the logistic regression algorithm from Christoforos et al. [43].
If the user makes a decision (i.e. by pressing a button) in a Flanker Task at time \( t \) we can create an epoch \( x(t) \) for a channel around that decision moment. The total length of \( x(t) \) is \( l \) samples (which is the number of samples in that epoch). We need to design a supervised classifier so that its output:

\[
y(t) = \beta^T * x(t)
\]

is expected to be maximally discriminated between two cases: either there is ERN or no ERN in that time window. This output \( y(t) \) is a real number.

\( \beta \): coefficient matrix \([l \times 1]\) that is unique for each channel.

We tested this algorithm with different signal pre-processing methods. For example, we tested with \( x(t) \) as EEG signals from all channels at a time moment \( t \) following the method from [142]. This type of input does not give us a good classification. We also tried to design a coefficient matrix for all channels but did not receive a good result.

In our approach \( x(t) \) is signal samples of a time windows (combination of two windows: before and after pressing button). Figure 33 presents the procedure of classifying an EEG epoch which is one of two types: ERN and no ERN.

**Figure 33. Online Single Trial ERN Detection Procedure**

The next step is to find the coefficient matrix to satisfy the expectation of \( y(t) \).

**Figure 34. Coefficient matrix extractor**

Figure 34 explains the procedure of finding the coefficient matrix \( \beta \) using a supervised logistic regression machine learning technique which is briefly explained in [43]. Once the coefficient
matrix is found, it can be used to classify an EEG epoch into one of two: there is ERN or there is no ERN.

5.4. Experiment 1: Flanker Task

To validate the above algorithm on a low cost, portable EEG device, we designed a Flanker Task that is similar to [142]. Twelve local students (7 males) were recruited to participate in this study. Each completed this experiment individually.

To mimic the experimental conditions in [47, 195] for a typical Flanker task experiment and reduce noise in data collection, the participant was seated in front of a screen in a dimly lit room. Participants were told to sit comfortably and minimize eye movement and blink as infrequently as possible while performing the task.

Participants were asked to perform a version of Flanker Task where they had to press one of two keys to specify the direction of a central arrow that was bounded by flanker arrows. There were two types of arrows, each type had two stimuli: congruent stimuli (<<<<<< and >>>>>>) and the incongruent stimuli (<<><< and >>>>>). All 4 stimuli were used in our trials in a random order.

For each trial there was a fixation cross in the centre of the screen for 500ms. It was replaced with one of four stimuli. The stimulus was presented for 100ms before the screen was cleared. Participants were asked to respond by pressing one of two buttons corresponding to the direction of the central arrow. At this time a string “-” appeared to mark each interval of 1000ms waiting. After the participant’s response, the screen remains clear for 500ms. All stimuli were presented in white font on a black background. At a viewing distance of around 100 cm, the visual angle of the arrow stimuli was 0.4° vertically and 0.6° horizontally, and between them was 0.3° space. This procedure is similar to the task carried out in [47] and [61]

Participants wore an Emotiv EEG neuroheadset that has the ability to capture raw EEG in 14 channels (of the 10-20 international system) from different locations around the human head. At first, participants were given one practice block of 40 trials. After that, they performed 4 blocks of 40 trials in which EEG signals were collected. Participants had 2 minutes to rest after each block. We collected a total of 1920 trials from all participants of the experiment.

5.4.1. Signal Processing

The EEG signals, captured from the Emotiv headset, were divided into 2-second length epochs (1 second before and 1 second after the key press moment). With the sampling frequency at
128Hz (after downgrading from 1024Hz), the length of each epoch is 256 samples. The first 200ms of each epoch were used to remove DC offset following which all epochs were filtered in 1-10Hz to remove components that are not in the ERN frequency bands of that particular epoch. Figure 35 shows examples of single trial ERN in the two cases when the trial was incorrect (Fig 4a) and correct (Fig 4b). These examples were picked to demonstrate the difference in the ERN pattern from a single trial.

Figure 35. Examples of single trial ERN for two cases: incorrect (a) and correct (b). The x axis is number of samples (1 sec = 128 samples). The y axis is amplitude (µV). The red line is the key press moment (sample number 127).

5.4.2. Results

For each channel, half of the trials (= 80 trials) were used for training via a logistic regression technique and other half were used for testing. We performed a t-test on $y(t)$ (calculated by multiplying $x(t)$ with the coefficient matrix) for each channel per user to check if there was a significant difference between two types of output (ERN and no ERN epochs). Based on the results of the t-test we found that on average F3, F4, F8, FC5, FC6, AF4 were the channels that yielded significant differences in $y(t)$ implying they are best suited to detect an ERN pattern from a signal $x(t)$.

From the test trials we found that using data from F4 channel can discriminate the two types of responses best: 69.7% of correct trials were classified as correct and about 70.3% of erroneous trials were classified as incorrect.

Figure 35 shows the average accuracy of different sensing channels on which the algorithm works most effectively. It was observed that those channels, which correspond to the frontal lobe, produce better classifications than other channels. This matched the literature about the origin of ERN, which is from the Anterior Cingulate Cortex (ACC) [196]. This also satisfies the fact that ERN has the frontal-central distribution of the human brain.
Figure 36. Classifier accuracy on a single trial basis

Figure 37 illustrates the average EEG signals over all epochs belonging to two cases: correct and incorrect.

Figure 37. EEG signals at channel AF4 with number of samples in the x-axis (1 sec = 128 samples) and uV in the y-axis. (a) Correct Epochs and (b) Incorrect Epochs. The vertical red line is the moment when key was pressed.

Figure 38. ROC curve for 4 channels: AF4, F3, F8, and FC6
To investigate the efficiency of the classifier further, we used the Receiver Operating Characteristic (ROC) Analysis [181]. It has two distinct inputs: hit rate (or true positive rate) and false alarm rate (or false positive rate) as two separate performance measures. ROC analysis has been used in machine learning recently to justify how good a classifier is by evaluating its discriminating power [91]. Figure 38 shows four ROC curves of 4 frontal-central channels. The further the curve is from the diagonal line, the more effective that classifier is. The area under the curve (AUC) gives an indication of how well the classifier is performing. The AUC of 1 indicates a perfect classifier and 0.5 indicates a random chance of classification. Our classifier achieves an average of 0.77 across all channels. The AUC for the best six channels (F3, F4, F8, FC5, FC6, and AF4) are shown in Figure 39.

![Figure 39. AUC for different channels](image)

### 5.4.3. Discussion

The result of the above Flanker Task demonstrate that Emotiv device is capable of capturing EEG signals with sufficient quality for a classifier to be able to detect ERN with an accuracy of about 70%. This initial result suggests that we can achieve ERN detection rates that can benefit interactive applications where these can be further improved through manipulating the feedback mechanism such that any likely error (or mistake) is noticed earlier by the user. This can provoke higher amplitude of ERN, further increasing the probability of accurately detecting it. As a result it will enhance the advantage of using ERN as the system can “sense” that the user made a mistake and interrupt or undo the last action.

### 5.5. Experiment 2: Buttons selection

Compared to Flanker tasks, a visually rich environment, like the one a user might encounter in an interactive application requires more users' mental workload. This will trigger neuronal activities from different parts of the brain; all of which can interfere with ERN signals. Mouse inputs in interactive tasks require free hand movements over longer distance which can cause bioelectrical signals from hand muscle activity. These activities will bring artifacts to EEG
signals [169]. For these reasons it is not clear whether ERNs can be detected in a visually rich environment.

The main goal of this experiment is two-fold: first to confirm that an interactive task other than Flanker Task will produce an ERN pattern that can be measured using our portable EEG set and secondly to check whether the pattern can be classified using our online detection algorithm.

We designed a task that retains the main elements of a multiple-choice RT task while at the same time offering a visually rich environment that goes beyond the Flanker task and be closer to what might be expected of an interactive application.

5.5.1. Task: Button selection

![User interface of the task at the moment of (a) start trial and (b) cursor is near the desired button.](image)

In this task the user had to select an object with the size of a 160 x 55 pixel button. In each trial, the user was asked to select a button in a limited time. There were 7 buttons with text from 'Link 01' to 'Link 07' that were arranged in an ascending order from top to bottom. The trial began when user clicked on the START button. A message at the top of the screen showed which button needed to be clicked (Figure 40a). The user then moved the mouse cursor toward the required button. A timer was placed at the top right of the screen to show elapsed time since the trial started. The timer was to put pressure on user to complete the trial as quickly as possible. The user must finish the trial in a limited time (1.4s) otherwise a TIME OUT message would appear and the trial restarted. We chose 1.4s because 1.3s was the average time for an experienced user to complete the task. When the mouse cursor was moving toward the desired button and there was 30 pixels distance remaining, the order of the buttons may change (Figure
40b). The probability for the buttons' order to change was 50%. Because the position of the desired button was changed when the cursor was very near, user may not have enough time to change their decision to click on the intended button but still aware of the result of their action. The intention was to provoke an ERN in this case. When user clicked on a button, a message notified user if they had clicked on the correct one. There was a 3 second waiting time before a new trial started.

5.5.2. Design and procedure

Nine students (6 males) recruited from the local university took part in this experiment. All of them knew about the Emotiv device but had never tried it before. Each of them was given the consent form and information sheet informing them about the task details. They were asked to sit comfortably but alerted to the task. They were also instructed to blink less during the experiment, especially around the deciding moment (mouse click). The experiment took place in an office environment with ambient noise.

Participants were given some practice trials to become familiar with the task until they confirmed they were ready to start the experiment in which data was collected. Each participant performed 4 blocks of 40 trials each and there was 2 minute break between each block.

5.5.3. Data Collection and Analysis

Signals were collected using the Emotiv neuroheadset. They were then divided into 1.5 second epochs around mouse click moments, 0.5 seconds before and 1 second after. The moments that a participant saw the result of their action were also captured as it is assumed that was when they realized that they have made a correct or incorrect choice. The epochs were then divided into two groups where the first group was used to compute the coefficient matrix ($\beta$). The second group was formed with the remaining data.

This coefficient matrix was then used to classify epochs of the second half. There were two cases to classify: participants clicked on the desired button successfully or they performed it unsuccessfully. The classifier performed the classifying task on a single trial basis. Epochs were divided based on the probabilities of having ERN and not having ERN. The result then was compared with the ground truth of participants' confirmation.

We do the classification and analysis offline but using the online detection algorithm. The primary goal here is to verify the effectiveness of the algorithm not to provide users with real-time feedback in an interactive task.
5.5.4. Result

We visually inspected the EEG signals around the moments when a participant clicked a mouse button to see if ERN patterns can be observed or not. ERN pattern was clearly visible in epochs where participants made an incorrect decision (clicked on a wrong button). As the trial result was displayed immediately after user clicked, the epochs of displaying moments were the same as the epochs around mouse click moments.

Figure 41 shows the mean of EEG signals around the moments when a participant click (select an object).

![Figure 41. EEG signals around moments where a participant made an incorrect decision (a, c) and correct decision (b, d) at channel F3 (top) and FC5 (bottom). The red lines mark the moment that the mouse button was clicked.](image)

In addition, the Logistic Regression algorithm can classify epochs belonging to those two cases: correct and incorrect responses. The coefficient matrices ($\beta$) of the same participant from the first half of data were multiplied with the extracted epochs $x(t)$ to produce output $y(t)$. We perform a t-test on $y(t)$ for each channel per user to check if there is a significant difference between two types of output: correct (user clicked on a correct button) and incorrect (user clicked on an incorrect button). We investigated further based on the t-test result and found that the accuracy of the classifier on incorrect decision moments and correct decision moments were 64% and 67% at channel F3 and 63% and 69% at channel FC5 (Figure 42). The results from this experiment are similar in accuracy to the result of the Flanker Task. It is notable that those two channels that have the best classifying accuracy are in frontal-central part of human brain which is consistent with the result obtained from Flanker Task.
5.5.5. Discussion

The result obtained from this experiment confirms the existence of an ERN pattern in a non-Flanker task. Those patterns can be seen clearly in epochs when participants have made an incorrect decision. Moreover, this result confirms that ERN patterns appeared in a normal application in a working condition as long as the ERN triggering condition is designed properly.

The method investigated in this experiment demonstrates the ability to detect an ERN pattern using a off the shelf EEG headset. It also demonstrates the ERN detection on a single trial basis. This is promising for HCI designers as the ultimate purpose is not to develop an effective and robust classifier but to harness this type of ERP into HCI.

While one might argue that we processed EEG data offline, it is worth noting that we used a detection algorithm that is essentially an online detection algorithm that works on a single trial basis. This algorithm can be used as a light weight module and can be run in the background of the system. Thus, it can be easily brought to an online version with the same design and signal processing methods.

The questions that remains is how useful this classification level is when ERN is applied to interactive applications and is 65% to 70% accuracy enough in effectively assisting users. The answer varies depending on the types and purposes of each task. Therefore, we designed the next experiment in order to demonstrate at least one set of applications that benefit from this level of ERN classification accuracy.

5.6. Experiment 3

The goal of the experiment was to determine whether the accuracies of the ERN classifier could benefit interactive applications and improve users' performance. It is unlikely that ERN classifying rate can reach 100% at all time. Therefore we design this experiment to justify the trade-off between ERN accuracies and benefits toward users.
The chosen application for this experiment was similar to Superflick [154]. Superflick is a pointing technique that is based on Flick (sliding/throwing the object across the table) but adds a correction step. Flick is an open-loop technique, providing fast movement but requires practice to achieve accuracy. Superflick offers a "remote drag-and-drop" correction phase if the object is off the target. In our experiment we integrate a “simulated ERN detection” module into Superflick to test if there are any performance differences between detection accuracies.

There are several reasons why we pick this specific pointing technique to explore performance. The main motivation is that it is easy to establish ground truth to compare the different techniques. Secondly, Flick is a popular interaction technique that has been studied extensively in the HCI literature for both handheld and tabletop environments. Thirdly, the authors have experience with the Flick technique and understand how the Superflick design can be improved through the introduction of an ERN detection module. This allows us to examine the effect of various ERN detection rates on users' performance.

The simulated ERN detection module consists of a random function used to mimic the ERN classifier. There are 4 types of accuracy: 50%, 65%, 80% and 100% to simulate the probabilities that the classifier will detect ERN patterns. The random function was controlled so that accuracies were as precise as required. During the experiment, participants were not aware that the ERN module was simulated.

**5.6.1. Task and Technique**

Participants had to move an object onto a target. The task started with the animated ball to move (the main ball – the lower yellow circle in Figure 32) located at the middle bottom of the screen. The target was a green circle and was assigned randomly among 15 targets.

The participant used a stylus to flick the object onto the target. As soon as the flick action was completed (by lifting the stylus of the touch surface) a circle appeared at the estimated final position of the main ball (the estimated ball – the higher yellow circle in Figure 32) that was calculated based on the distance and duration of the flick gesture. This instant visual feedback allowed the user to know whether or not the ball would actually hit the target. This was done to help with potentially triggering an ERN (although we do not capture these signals but rely on our controlled random function). At this point the mimicked ERN detection module will be triggered causing either the red (mimicking ERN detected) or the green (mimicking ERN not detected) light to be visible. The main ball then moved from the start point toward the estimated ball with the speed of 1200 pixels per second. When the main ball was moving, the user had a
chance to remotely drag-and-drop the estimated ball to the target (as in Superflick). During the drag process, the main ball automatically moves toward the estimated ball. The trial finished when the main and the estimated balls met each other for over 200ms. This threshold was set to prevent accidental overlaps of the two balls. If they met inside the target, the trial was marked as successful and unsuccessful otherwise. A new trial started after 3s waiting and the user interface was reset to the original condition (with a new target).

The simulated ERN detection worked as follow: 250ms after the main ball left the start position, the ERN detection light, which was placed at the bottom right corner of the screen, started showing red if ERN pattern was detected (estimated ball was not inside the target) and green if not. In case of the red light, the main ball's speed decreased dramatically to 15% of the original speed (180 pixels per second). The user then had more time to remotely correct the position.

The task was built in C# for the experiment, and was installed in a bottom-projected tabletop system. The table was 105x88x106cm (w x d x h) with display size of 72.5x60cm and projector resolution of 1280x1024 pixels. Participants used a Wacom CTE 430 tablet (dimensions: 210 x 208, active area: 127.6 x 92.8mm) for input.

5.6.2. Design and Procedure

Nine local students (5 females) with age from 21 to 41 participated in this experiment. All of them had heard about the Emotiv headset but never had used it before. They were given a consent form and information sheet explaining about the purpose of the experiment. Participants were not aware that we were simulating ERN detection. They wore the EEG headset, stood comfortably but alerted to the task in front of the tabletop, and were led to believe that their EEG signals were used to assist them in their task. But instead of detecting ERN based on the measured EEG signals, we created a detection module which is a controlled random function. The accuracy of this detection module was controlled to simulate the accuracy rates.

Participants were given 1 block of 50 trials or more to practice until they were familiar with the task and ready to start. After that, they performed 4 blocks (with minimum 50 trials) with ERN accuracies of 50%, 65%, 80% and 100%. A block was finished when number of trials was more than the minimum number of trials and ERN detection accuracy reached the required number.

5.6.3. Result

Figure 43 illustrates the percentage of successful trials (the main ball ended inside the target). We applied one-way ANOVA test with Bonferroni post hoc multiple comparisons on the data.
and found that there was significant difference between groups ($F = 17.517, p < 0.001$). However, there were no significant difference between groups of 65%, 80% and 100% ($p > 0.5$).

We also analysed the number of unsuccessful trials when there were red and green lights separately. The intention was to check if users made mistakes naturally when the main ball moved slowly (red light); and if there were more unsuccessful trials when the main ball moved with full speed (green light).

![Figure 43. Accuracies of Superflick with different ERN detection accuracies](image)

Their mean values are shown in Table 2. We ran one-way ANOVA test with Bonferroni post hoc multiple comparisons on those data and found no significant difference between groups in case of Red light ($p > 0.5$). It was noteworthy that even there was significant difference between groups in case of Green light ($F = 11.741, p < 0.01$), there was no significant difference between groups (50%, 65%), (65%, 80%), and (80%, 100%) ($p > 0.5$).

<table>
<thead>
<tr>
<th></th>
<th>50</th>
<th>65</th>
<th>80</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Red</strong></td>
<td>0.12</td>
<td>0.03</td>
<td>0.06</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>Green</strong></td>
<td>0.47</td>
<td>0.35</td>
<td>0.22</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 3. Mean % of unsuccessful trials with the two indication light

**5.6.4. Discussion**

The results obtained from participants show that when integrate ERN detection module with 65% accuracy was as good as 80% accuracy. This was proved as no significant differences were found between two groups in term of overall accuracy and when ERN was not detected (incorrectly). Our results also show that integrating an ERN detection module might not prove beneficial if the detection accuracies are lower than 65%.

The classifier with 65% accuracy may provide higher error rate compared to 80% and 100% accuracies but still benefits interactive tasks. If a system has very high accuracy (less error rate),
it can promote hasty commitment to selections [83]. This is because users overly rely on the system and know that there is little cost for making mistakes. In the real world, the cost of an error may be much higher. Therefore, a system with lower classification accuracy may require more attention but has lower cost of recovery hence still offers benefits to the users.

These results confirm that ERN classification accuracy with Emotiv can benefit interactive tasks as good as other expensive devices, yet offer the advantages of portability, low-cost and instantaneous classification.

This experiment also shows an opportunity to assist users in pointing and table top applications using ERN. In these tasks, objects are usually out of range for users to select therefore they need to use several techniques such as drag-and-drop, radar, etc. ERN integration does not replace these techniques but provides better performance and more precise selections for users.

5.7. Discussion

ERN has the potential to enrich interactive applications. The results from the studies in this chapter can provide guidance on how best to begin harnessing ERN for such interactive experiences.

5.7.1. Employ ERN to Assist Interactive Tasks

The studies in this chapter have shown that it is possible to detect ERN patterns using an off-the-shelf EEG headset on an online single-trial basis. If integrated into interactive tasks, an ERN detection module provides another medium for HCI designers to access users' intentions, which is intuitive and directly from the users' brain. This module can be designed as a lightweight background feature.

An ERN pattern will appear in any multiple-choice RT task, of which the button selection task used in this work is just one example. The ERN pattern can be detected within 150ms of the event onset. This means within this time window the user becomes aware that they have made a mistake; the interactive application can know this and respond. In most cases this type of information is not available without access to the user’s EEG signals. Even if knowledge of user error is available in some other way (such as text auto correction, prediction techniques) it may not offer as fast response as ERN detection.
The nature and scope of usage of the ERN signal will depend on the usage context, the creativity of the designer and on the ability of the user to ignore recommendations that are incorrect or inappropriate.

5.7.2. Dealing with incorrect classification

Detecting ERN correctly is a challenge. We can be reasonably sure that even the best classifier will never achieve 100% accuracy in an online single trial system. Consequently, an interruption management system is needed to be integrated into the system. This can reduce the disturbance to users when the system makes suggestions based on incorrectly detected ERNs. For example, if a user chooses to ignore a suggestion that pops up based on a confirmed ERN, the pop up must not prevent the user interacting with the user interface. Designers wishing to use ERN in their applications must be careful to ensure that such ERN-based suggestions do not stop the user from working with the system.

5.8. Summary

This chapter offers some valuable guidelines for HCI designers through the experiments’ results. We show that ERN patterns can be detected using an off the shelf EEG headset on an online single trial basis. Moreover, we apply this model to an interactive task to illustrate that it can work with normal interactive applications in a working environment with ambient noise. We finally show that accuracies in the region of 65 to 80% are sufficient for ERN to be effectively integrated into HCI applications. In our discussions we suggested novel ways in which HCI applications could benefit from ERN.

This chapter explored ERN’s usability for interactive tasks. ERN is detected when a person make a mistakes in the traditional Flanker tasks and a common interactive applications. However, the investigation in this chapter focuses on the person who performs the tasks (or the executer/ performer). There is also another scenario in HCI where people work in a collaborative setting where it is easy to spot another people committed or going-to-be mistakes. We will investigate this in the next chapter by harnessing ERN in observing interactive tasks.
Chapter 6. Error Related Negativity in Collaborative Applications

In this chapter we continue the investigation of employing ERN for a passive BCI-based system. We demonstrated in the previous chapter that Error Related Negativity is triggered and can be detected on a single trial basis when a user either makes a mistake or the application behaves differently from their expectation. It can also appear while observing another user making a mistake. This chapter\(^4\) investigates ERN in collaborative settings where observing another user (the executer) perform a task is typical and then explores its applicability to HCI. We first show that ERN can be detected on signals captured by commodity EEG headsets like an Emotiv headset when observing another person perform a typical multiple-choice reaction time task. We then investigate the anticipation effects by detecting ERN in the time interval when an executer is reaching towards an answer. We show that we can detect this signal with both a clinical EEG device and with an Emotiv headset. Our results show that online single trial detection is possible using both headsets during tasks that are typical of collaborative interactive applications. However there is a trade-off between the detection speed and the quality/prices of the headsets. Based on the results, we discuss and present several HCI scenarios for use of ERN in observing tasks and collaborative settings.

6.1. Introduction

To date most examples of ERN have focused on detecting the signal by recording signals from the executer of the task. Recent endeavours [50, 132, 182] suggest that ERN signals even appear while observing another user making errors. These studies confirm the existence of negative potential within 250ms after the event onset and analysis of the origin of the signal confirm that they are ERN signals. In these studies the observer did not have much opportunity to anticipate the actions of the executer and often relied on the answer displayed. In HCI scenarios, the executer’s actions may become visible to the observer even before the action is committed. For example, in collaborative tabletops, an observer has awareness of an executer’s actions through their reaching gesture [147].

\(^4\) Work described in this chapter was published in CHI ’14 as “Error Related Negativity in Observing Interactive Tasks”.

89
We investigate the effect of anticipation in an observing task where the outcome of an action is revealed before that action is committed which has not been investigated with ERN in an observing task before. This can have many applications in pair-programming, collaborative tabletops and emergency response applications. We first investigate these anticipation effects with an expensive clinical EEG system to demonstrate that the signal is present and can be detected. Following this we investigate if an Emotiv headset has the similar detection capability.

We start by repeating an experiment from van Schie et al. [182] with an Emotiv headset to demonstrate that Observer ERN can be detected in a single trial and online basis, with the accuracy up to 64%, using a commodity headset. After that, we investigated the anticipation effect in observing tasks when one person observes another person committing errors. Our results show that there are ERN-like patterns detected in the observer’s EEG about 368ms after the initial movements happened and 55ms before the errors are committed. Following this result, we then show that the Emotiv headset can capture these patterns in the same experiment settings. Finally, we discuss the implications of our results on interactive applications.

The contributions of this chapter are: (a) we demonstrate that off-the-shelf EEG devices like an Emotiv EEG headset can capture ERN in an observing task from channels in the frontal-central part of the brain; (b) we investigate the anticipation effects in collaborative settings demonstrated by the ERN detected in an observer’s EEG signals before the action is committed; (c) through a final experiment we show that these anticipation effects can still be demonstrated using off-the-shelf EEG devices such as the Emotiv EEG headset.

6.2. Related Work

6.2.1. Error Related Negativity (ERN)
ERN is a form of an ERP triggered in the brain when a person makes a mistake or the application behaves differently from their expectation [47]. This pattern is produced in a person’s brain when they are aware of the obvious error(s) that they have made; either through system feedback or individual realization [68]. ERN peaks within 150ms after the action onset and has amplitude varied in accordance with the awareness of the mistake. Interestingly, ERN also appears when the user is confused about their last decision [47]. This pattern has been discovered by researchers using both expensive devices [47, 195] and commodity devices (such as Emotiv) [184]. ERN is useful in interactive applications when it is detected immediately following the triggered moment. To this end, researchers have looked at detecting this pattern in real-time, and on a single trial basis with accuracy up to 80% [43, 68, 142].
[184] showed that this detection can be done using a low-cost and off-the-shelf headset like Emotiv with around 65% accuracy. This makes ERN more accessible to game developers and other consumer application designers.

**6.2.2. Error Related Negativity (ERN) in Observing Task**

Recent endeavours [22, 50, 132, 182] suggest that ERN signals even appear while observing another user making errors. For example, Miltner et al. [132] reported the generation of an ERN potential when participants observed errors committed in a choice reaction time (RT) task. Moreover, van Schie et al. [182] reported an ERN pattern elicited in an observer when observing another person performing a modified Eriksen flanker task [60]. Other researches have confirmed these results by showing observed ERN following observations of other’s errors [22, 132]. These results suggested that similar neural processes trigger the detection of a person’s own error as well as the detection of others. Interestingly, de Bruijn et al. [50] aimed to disentangle the dependency of ERN on error or reward. Their results show that the performance monitoring as reflected by the ERN is error-specific and not directly dependent on reward. As a result, our study focuses on the error awareness of the observers.

To date, all studies of ERN in observation tasks only look at situations where the outcome of the action becomes visible around the moment where the action is committed. Furthermore, these studies focus on detecting the ERN pattern using averaging methods for the purpose of confirming the pattern’s existence. We are not aware of studies that investigate the context where the observer can anticipate the executer’s actions through their gestures towards the target.

In these contexts, anticipation effects can cause the observer to trigger an ERN before the moment when the executer finishes the action. Successfully detecting this pattern can be useful in many HCI applications such as collaborative tabletops, pair-programming and emergency response scenarios.

**6.3. Experiment 1: Eriksen Flanker Task on PC**

The purpose of this experiment is twofold: first to validate, based on the study in [182] and [50] that we can detect the ERN signal on an observer using off the shelf brain sensing technology; second to verify that the ERN pattern can be detected using the classifier presented in [184] using an off the shelf headset.
6.3.1. Task and Procedure

Each participant performed the experiment paired with an actor. We used an actor in all of our experiments to trigger the ERN signal on the participant. Only the participants wore an Emotiv EEG headset that has the ability to capture EEG signals of 14 channels in the 10-20 international system. They were told about the experimental goals to explain why they had to wear the EEG headset and not the actor. Participants were not aware of the actor's role and were led to believe that the actor is another participant. As the window of recognizing the ERN signal is very small, the role of the actor is to perform certain actions that will maximize the ERN detection from the observer during the task period.

The experimental procedure is similar to the task carried out in [47] and [61]. To mimic the experimental conditions for a typical Flanker task experiment and reduce noise in data collection (such as in [47, 195]), both the actor and observer were seated side by side in front of a screen in a dimly lit room. Participants were told to sit comfortably and minimize body, facial, and eye movement as well as to blink as infrequently as possible during the task.

**Trial:** Each trial began with a black screen for 3s, followed by a fixation dot in the center of the screen for 200ms. After that, the screen remains clear for 200ms before one of four stimuli was displayed for 300ms. There were two types of arrows, each type had two stimuli: congruent stimuli (<<<< and >>>>) and incongruent stimuli (<<< and >>>). All four stimuli were used in our trials in random order. At a viewing distance of about 100cm, the visual angle of the arrow stimuli was 0.4° vertically and 0.6° horizontally, and between them was 0.3° space. The executer was asked to press the direction key as soon as they saw the stimulus to indicate the direction of the middle arrow. The input device used for this key press was a remote control (Genius Media Pointer). The pressed key (< or >) was displayed for 1s afterward. Note that the executer can be either the participant or the actor depending on the trial condition, as explained below.

The observer was then asked to rate the correctness of their last action by choosing one of three options: 1. Sure correct, 2. Do Not Know, and 3. Sure Incorrect. This was input through a mini keypad.

**Practice trials:** Participants were asked to perform a version of the Flanker task where they had to press one of two keys to specify the direction of a central arrow that was bounded by flanker arrows. This was done to make the participant familiar with the task that will be performed by the actor giving them an idea of what to expect as quick as possible when the stimuli appears.
Each participant performed a practice block of 40 trials where they had to press the direction key and rate the correctness of their action in each trial. During this time, the actor sat beside the participant and played a dormant role. In these trials, the directional arrows, which appeared on the screen after participants pressed a button, were exactly like the pressed button. We did not collect EEG signals or any data during practice block.

Figure 44. Observation of Flanker Task with the executer (left) and the observer (right)

**Experimental Trial:** After finishing the practice block, the participant and the actor switched places. The actor now had to press the directional keys and the participant (the observer) rated the actor's last performance. However, the actor hid her hands under a box to prevent the observer from seeing which button was physically pressed (see Figure 44). The observer was instructed to focus on the screen to judge the correctness of the actor's answers. As in [22, 132] participants were instructed to silently count the number of incorrect selection. However, the result displayed on the screen was independent of the button press: 40% of the time the computer displayed a wrong answer (the 'pressed button' is opposite of the middle arrow of the stimulus). This allowed us precise control of the experiment settings.

The actor performed 4 blocks of 50 trials each with 3 minutes break after each block. We collected 200 trials from each participants of the experiment. We recruited 6 university students, aged between 21 and 29 years old to give us a total of 1200 trials for this experiment.

**6.3.2. Data Collection and Analysis**

The EEG signals, captured from the Emotiv headset, were divided into 2s length epochs; 1s before and 1s after the key press moment. With the sampling frequency at 128Hz the length of each epoch is 256 samples. The first 200ms (about 25 samples) of each epoch were used to
remove DC offset following which all epochs were filtered in 3-8Hz to remove components that are not in the ERN frequency bands of that particular epoch.

6.3.3. Result

We employed the classifying method briefly described in [43] and validated in [142, 184] to analyse the data. It can be summarized as follow.

For each channel, half of the trials (= 100 trials) were used for training via a logistic regression technique and other half were used for testing. We performed a t-test on the classifier output for each channel per user to check if there was a significant difference between two types of output (correct and incorrect epochs). Our results reveal significant differences in classification rates (for correct and incorrect epochs) for FC5, F7, and F3 channels. The classification rate for the rest of the channels was not significantly better than chance.

Figure 45 (left) shows the average accuracy of three sensing channels. These three channels correspond to the frontal lobe which is in line with the literature about the origin of ERN (the Anterior Cingulate Cortex - ACC) [196]. Figure 46 illustrates the average EEG signals at F3 and F7 over all epochs belonging to two cases: correct and incorrect. The ERN peaks at about 250ms after the event onset (key pressed and result displayed on the screen).

Figure 46. Average EEG signals at channel F3 (top) and F7 (bottom).
We did a Receiver Operating Characteristic (ROC) Analysis to investigate the efficiency of the classifier [181] by evaluating its discriminating power [15]. ROC analysis uses two distinct inputs: hit rate (or true positive rate) and false alarm rate (or false positive rate) as two separate performance measures. Figure 47 shows ROC curves for F3 and F7 channels. The further the curve is from the diagonal line, the more effective the classifier. The area under the curve (AUC) gives an indication of the performance of the classifier. An AUC of 1 indicates a perfect classifier and 0.5 indicates a random chance of classification. Our classifier achieves an averaged value of 0.66 over these three channels (F3, F7, and FC5). Their AUC are shown in Figure 45 (right).

Figure 47. ROC curves for channels F3 and F7

6.3.4. Discussion

This study demonstrates that Emotiv EEG headset is capable of capturing EEG signals with sufficient quality for a classifier to be able to detect ERN pattern with an accuracy of about 65%. The characteristic of detected ERN patterns is similar to previous studies using more expensive devices [182]. Our results further demonstrate that the classifier described in [43] and validated in [142, 184] can be used to detect ERN in observer tasks. This means an interactive application can have a similar classifier detecting both observer and executer ERNs with minimal modification to the software.

One example of interactive tasks that can benefit from our results is in pair programming where two programmers work together as partner, on the same machine, to complete a programming work with their roles switched frequently [190]. One programmer is the driver, who performs all “on computer” tasks. The other is the observer or navigator who reviews each line of code and points out the errors as it is being written [127]. Usually this error correction process requires the observer to interrupt the programming process, point out the location of the errors in his opinion either using his hand or using the mouse/keyboard that the driver is controlling.
However, our experiment results suggest that monitoring the appearance of ERN in the observer’s EEG can speed up this error correction process. Each action of the driver is considered a trial of a multiple choice RT task. If the ERN is detected in the observer’s brain, it can be used to pinpoint the whereabouts of the error (e.g. highlighting the code section) which is 250ms after the action’s committed moment. The system can also provide suitable suggestions based on the context of the location where the error was triggered.

In addition, we believe the detection accuracy can be enhanced by improving the visibility and awareness of the performer’s actions through better visualization. This will elicit higher amplitude of ERN that can improve the classification rate. More detailed implications of this study to HCI are discussed at the end of this chapter.

6.4. Observer ERN in a Tabletop Task

In the first study the participant did not have any opportunity to see the actions of the actor and often relied on the answer displayed. In collaborative HCI scenarios like around a tabletop, a participant can usually anticipate the executer’s actions through their gestures which reveal the outcome of an action before that action is committed.

The awareness caused by anticipation gives the observer more time to form an opinion on the action. This could potentially reduce the time-critical aspect of the ERN leading to a low signal quality. Alternatively, the observer may reach an opinion of the executer’s action as soon as they see the initial cues. In this case, we may detect a good quality ERN in the observer well before the executer has even completed the action.

We are not aware of any experimental investigation of the effect of anticipation in an observing task where the outcome of an action is revealed before that action is committed. A primary objective of our second experiment is to examine this effect. If we can detect an ERN, our second objective is to determine the moment when the observer elicits an ERN.

6.4.1. Experiment Setup

Each experimental session involved two users – an executer and an observer seated around a rear-projected FTIR interactive table of height 76cm. The executer was an actor trained to do this study while the observer was our experiment participant. The executer and observer sat opposite to each other so that they were aware of each other’s movements and actions (similar to the setup in [182]). The projection area of the table was 72cm x 48cm (resolution 1024x768 pixels) and touch detection was done through a Point-Grey Dragonfly 2 camera.
All participants performed the task with the same actor (the person at the top in Figure 48). However, participants were led to believe that the actor is just another participant like himself or herself. In order to study the effect of anticipation, we had two experimental conditions – a close layout and far layout.

In the close layout the actor’s buttons (size 100x100 pixels each) were placed right under their hands and next to each other. The participant would struggle to see the actor’s hand movement limiting their ability to anticipate the outcome before the actor commits to the action. Any ERN elicited in the participant (observer) in this condition would be because of the committed action of the actor.

In the far layout the two confirmation buttons were placed close to the left and right edges of the table (200 pixel gaps to left and right edges with 524 pixels distance between them; 80 pixels gap from two buttons’ centers to the edge of the actor’s table side). The actor was asked to keep his hands touching two touch sensors placed on his side of the table (see the table’s top edge in Figure 48, right). In order to select a response the actor had to lift one hand from the rest position and move towards the button. This action took approximately 400ms giving the participant sufficient time to anticipate the actor’s actions. The actor’s other hand was left in the rest position touching the sensor during this action.

6.4.2. Task and Procedure

Each participant wore the EEG cap of the BeLight System from EbNeuro [2]. They observed and gave feedback about the correctness of the actor’s performance in a version of the Eriksen flanker task.

Each trial began with a black screen for 3s, followed by a fixation dot in the center of the screen for 200ms. After that, the screen remains clear for 200ms before one of four stimuli was displayed at the center of the table for 300ms. All four stimuli were used in a random order.

In both close and far layouts (Figure 48 left and right), the actor then touched one of two buttons to indicate the direction of the middle arrow. In close layout there is no noticeable hand movement whereas in the far layout the actor has to reach the button to make a selection.

After the answer button was selected, it was highlighted in red for 1s before turning back to the initial colour. This is to re-enforce to the observer the actor’s selection. The participant was then asked to rate the correctness of their last action by choosing one of three options: 1. Sure correct, 2. Do Not Know, and 3. Sure Incorrect. During the experiment, the observers were asked to
minimize their body and facial movements as well as blink as infrequent as possible. They were also instructed to guess the outcome of each answer as quick as possible although they could input that answer at their own pace. As in [22, 132] participants were instructed to silently count the number of incorrect selection. Since it was not possible to control the ratio of correct and incorrect trials to the extent done in Experiment 1, the actor was asked to keep the incorrect selection rate to between 25 and 30%. The actor received practice to minimize difference in action between correct and incorrect gestures.

![Tabletop Flanker task a) close, b) far layout.](image)

Each participant performed 2 blocks of 80 trials each per layout. The order of presentation of the blocks was controlled using a Latin square to reduce order effects. Participants received 3mins break between blocks and the whole experiment took about 90mins per participants including about 40mins of setup time.

Ten participants (7 males) between the age of 19 and 31 volunteered for the study. All were from the local university and did not participate in the earlier experiment. All participants had normal or correct-to-normal vision, and none of them were colour-blind. None of them had undergone brain surgery or had any known neurological disorders. Participants wore EEG cap during the experiment. They received a financial compensation for participation in the study.

### 6.4.3. Data Collection and Analysis

EEG signals were collected using the BeLight System from EbNeuro [2]. The sampling frequency was set at 4 kHz and electrode impedance was below 5 KΩ. EEG signals were then down sampled to 1 kHz before being divided into 2s epochs around the touching moment,
which is also the moment that the results were displayed or the touched button was highlighted in red. It is also assumed to be the moment when participants became aware that the actor had or had not made a mistake. Epochs were averaged using the first 250ms signals of each epoch. They were then passed through a 50Hz notch and [3-8Hz] band pass filters. These epochs were then divided into two groups where the first group was used to train the classifier (to calculate the coefficient matrix $\beta$). The second group was formed with the remaining data.

The coefficient matrix $\beta$ was then used to classify epochs of the second group. We performed this procedure separately for the two layouts. In each, there were two categories to classify: correct or incorrect. Epochs were divided as correct and incorrect trials based on the participants ratings. The classifier performed the classification on a single trial basis. The result then was compared with the ground truth of each trial.

6.4.4. Result

We collected 3200 trials, 1600 each for close and far layout. The number of incorrect trials was 27% in the close layout and 30% in the far layout. 7 trials in the close layout and 30 in the far layout were rated as “Do Not Know” by the participant.

![Figure 49. Averaged signals of the ‘close’ layout at Cz](image)

**a. Close Layout Analysis**

Figure 49 shows the averaged signals in Cz channel in the close layout. It can be seen that there is a difference, however small, between incorrect trials (red curve) and correct trials (blue curve). Latency of the highest peak after the touch button was pressed was 140ms. However, the topographic distribution (Figure 50) shows a clear difference between correct and incorrect trials around the time of the peak. It can be seen that there is a high density of negative EEG at the central and frontal area around the latency of the highest peak after the key press onset (Figure 50, middle).
Figure 50. Topographic distribution of correct trials (top) and incorrect trials (bottom) in the ‘close’ layout at the intervals: key press (left), 100ms after (middle), and 200ms after (right).

b. Far Layout Analysis

In the far layout, the differences can be seen clearer with the averaged EEG signals in Figure 51. The left vertical line depicts the averaged moment that the actor lifted up his hand and started to reach to the answer button. The right vertical line depicts the moment that the actor pressed the answer button. It took the actor averagely 428ms (correct trials) and 423ms (incorrect trials) to touch the answer. There was no significant difference found between these two durations (p > 0.05). The signals peak at 55ms before the answer buttons were touched. The topographic distribution (Figure 52) shows a clear difference (in the density of negative EEG at the frontal and central area) between correct and incorrect trials at the time of the peak (about 50ms before button touch, left figures) and at the time of the touch (middle figures).

Figure 51. Averaged signals of the ‘far’ layout at Cz

Using the same classifying method described in the previous experiment, we found that several channels in the frontal-central area of the brain yield successful classification. More details are in depicted in Figure 53. The highest classification rates are from channel Cz for both close layout (correct/incorrect = 70.43%/68.71%) and far layout (correct/incorrect = 73.02%/72.40%).
We also calculated the AUC for the above channels. Their values are plotted in Figure 11. Overall, the classifier achieved averaged AUC values of 0.6733 for the close layout and 0.6602 for the far layout.

![Figure 11](image1.jpg)

*Figure 52. Topographic distribution of correct trials (top) and incorrect trials (bottom) in the ‘far’ layout at the intervals: 50ms before key press (left), key press (middle), and 50ms after key press (right).*

![Figure 53](image2.jpg)

*Figure 53. Classification rates of close layout (left) and far layout (right).*

![Figure 54](image3.jpg)

*Figure 54. AUC for close (left) and far layouts (right)*

### 6.4.5. Discussion

The result of this experiment demonstrates that ERN can be detected in an observer in both hand layouts. It also shows that there is an anticipation effect where the ERN pattern in the observer’s mind appears about 55ms before the touching action is committed.
The experiment mimicked a classic Flanker task to keep our experimental settings in line with other research in this area. Interactive applications can trigger ERN using a similar paradigm. It may be possible to detect the ERN pattern earlier by making the actor’s movement clearer to the observers which can trigger the awareness sooner.

It is worth noting that the classifier used is light-weight module which can provide an output in less than 1ms for a 1kHz sample once the $\beta$ coefficient matrix is known for that participant (tested in Matlab R2013a 64 bit, Intel® Core™ i3 CPU 3.10GHz, 8GB RAM). In principle, this means that any interactive application could act on an executer’s action based on the ERN of an observer. For example, in collaborative tabletops a user could be asked for a stronger confirmation if the system detects an ERN in the observer. This could also benefit peer-learning activities where the observer and executer can constantly switch roles to learn from each other.

A drawback of this study is that the EEG cap used belongs to an expensive clinical EEG system. The cap is cumbersome to wear, tethered and does not lend itself to applications that require the observer to move freely within the experimental environment. This further limits the possibility of both the observer and the executer wearing the cap in a realistic environment. The following experiment aims to address this drawback.

### 6.5. Tabletop Observer ERN with Emotiv

The purpose of this experiment is to see if the Emotiv EEG headset can be used to detect the same pattern as in experiment 2. This would then open up the use of observer ERN to a wide range of interactive applications by detecting and predicting errors in observers’ mind.

#### 6.5.1. Task and Procedure

The experimental setup, task and procedure were identical to the previous experiment. We recruited 11 participants (8 males), aged between 20 and 31 years old. None of them took part in previous studies. The experiment took about 60mins per participant. This shorter experiment time, compared to the previous experiment, was due to the shorter setup time needed for Emotiv headset.

As with experiment 2, we collected 320 trials per participants yielding a total of 3520 trials from all participants of the experiment for both close and far layouts with 1760 trials each. The number of incorrect trials was 26% in the close layout and 28% in the far layout. 19 trials in the close layout and 34 in the far layout were rated as “Do Not Know” by the participant.
6.5.2. Data Collection and Analysis

Signals were collected at a sampling frequency of 128 Hz using the Emotiv EEG neuroheadset. They were then divided into 2s epochs around the touching moments, which is also the moment that the results were displayed. The analysis was the same as in previous experiment.

a. Result

Close layout: Figure 55 shows average signals at channel AF3 in the close layout over all participants and all trials. It can be seen that there is a difference between incorrect trials (red curve) and correct trials (blue curve) at the latency of 188ms after the moments where the button answers were pressed. This pattern has a negative peak appeared within the time period of 250ms after the event onset, similar to previous studies such as in [50, 182].

![Figure 55. Averaged signals of the close layout at AF3. The vertical line is the button touched moments](image)

Far layout: Figure 56 shows the averaged EEG signals for the far layout over all participants and all trials. There is also an ERN-like pattern which peaks at about 3ms after the answer button was touched. This indicates that if this is the ERN pattern, the observer must have the judgment about the actor’s action’s correctness about 250ms before which also is an indication that the observers have made the decisions before the answer buttons were touched.

Using the same classifying method described in the previous experiment on the collected signals from Emotiv device, we found that in close layout, channel AF3 yields the classifying results with 67.32% of correct trials were classified as correct and 64.57% of incorrect trials were classified as incorrect. Additionally, in far layout, channel F3 yields the successful classification with classification rates for correct and incorrect trials are 65.03% and 62.30% respectively (see Figure 57).
Figure 56. Average signals of the far layout at F3. The left vertical line is the averaged hand lift-off moments, the right vertical line is the button touched moments.

Figure 57. Classification rates of experiment 3

Additionally, we calculated the area under the curve (AUC) values for channels AF3 and F3 to justify how well the classifier performs on these two channels. Our obtained results show that AUC for channels AF3 and F3 were 0.6483 and 0.6486 respectively.

6.5.3. Discussion

The results from this experiment show that it is possible that ERN can be observed in collaborative tasks from the observers’ EEG. Although the channels that show the detection are limited to AF3 (close layout) and F3 (far layout), it is an indication that an off-the-shelf EEG headset such as Emotiv is capable of detecting this pattern. As a result, other commodity headsets which have access to these channels can also harness the advantage of detecting ERN pattern in observing tasks.

There is a shift in latency of the ERN peaks between experiments 2 and 3. The latency differences of about 40ms (close layout) and 60ms (far layout) are due to the devices used in each experiment; EbNeuro BeLight system (~USD 10k) and Emotiv headset (~USD 300). There are some postulates that may explain the differences between them:
• The buffer used in each device is different. While Emotiv waits for the buffer to be filled before pushing signals toward receiving device (Bluetooth dongle), EbNeuro pushes the signals continuously for each collected data sample. As the process of continuously pushing signals requires much more expensive Digital Signal Processing (DSP) module, this reflect the price difference between the two devices.

• Another possible reason is while EbNeuro push signals continuously at the original sampling frequency (4KHz), Emotiv samples EEG signals internally at 2KHz then downsamples the signals to 128Hz output. This down sampling procedure takes a time to process.

• Emotiv headset uses Bluetooth to transfer EEG signals to the receiving dongle. This process may take longer compared to wired transfer because it requires separate DSP module to encode and encrypt the signals before transmitting.

From the results of this experiment, it can be seen that there is a trade-off between how early the ERN pattern in observing tasks can be observed vs. the prices and quality of EEG headsets. Although the differences are small (from about 40ms to 60ms), they should be considered carefully to fit with the goals of each interactive task.

6.6. Summary

The experiments described in this chapter offer some valuable guidelines for HCI designers in order to employ ERN for a passive BCI system. We show that ERN patterns can be detected in an observing task using an off-the-shelf EEG headset on a single trial basis. Moreover, we show the anticipation effects in collaborating work where ERN can be detected before a committed action in the observer’s mind. We then extended our finding to a commodity headset to show that it can also detect the anticipation effect through the existence of ERN. HCI applications can benefit from our results in implementing ERN in collaborative and observing environments to pinpoint the position of an error. This may consequently prevent the mistake from committing or speed up the correction process. This is highly beneficial in some typical interactive tasks such as collaborative tabletop settings, war-room/ emergency response, and gaming scenarios. We will present these future directions in more detail in the next chapter.

6.6.1. Limitations of the classification method

Our classifier is based on a linear regression method described in [43] and validated for single trial ERN detection [142, 184]. Applying this method provides classifying rates of up to 73% accuracy which was shown to be beneficial to HCI [184]. However the classifier requires input
EEG signals ±1s around the time-locked event. This means that the classifier can detect the ERN pattern 1s after the action. However it can be used to pinpoint the moment of ERN which can be up to 50ms before the action. Despite this delay, this form of ERN detection can still be useful in many applications like in the pair programming case outlined earlier. A 1s delay will not affect the benefit of ERN in such activities especially when we can highlight the code-section that is in question. Furthermore, other classifier methods can improve accuracy and reduce the window size of input signals leading to earlier detection of the ERN after it is triggered. For example, a BCI competition on customized classifier increased P300 classification rates up to 96.5% and motor imaginary classification rates up to 94.2% [32].

Through the four chapters, we have investigated the usability of task engagement and ERN as well as capabilities of commodity EEG headsets to benefits EEG-based interactive applications. Our investigations were for both active- and passive BCI systems. In the next chapter, we will overview our findings, discuss the implications of the results on HCI, and examine the future directions of the works in this thesis.
Chapter 7. Discussion and conclusions

In the previous chapter, we presented the active and passive usage of commodity EEG devices to benefit interactive tasks. We provided guidelines in using task engagement for continuous application control as an example of active BCI, as well as an evaluation tool as an example of passive BCI. On the other hand, guidelines of how to harness ERN for interactive tasks were also given for demonstrating its use in passive BCI. In this chapter, we will look back as well as forward into the use of BCI with commodity EEG headsets for interactive applications.

7.1. Looking back

7.1.1. Contributions

The user studies conducted during this thesis aim at demonstrating the opportunities of using EEG signal to benefit everyday users. In these studies, we have gone through the process of capturing EEG signals, calculating task engagement, detecting ERN, and presenting several guidelines in harnessing their benefits. We have also provided guidelines on how to use captured EEG signals with commodity headsets in both active and passive BCI systems. Consequently, it is now possible to look back and summarize the research outcomes of this thesis as follows.

Chapter 3: Quantifying Task Engagement for Gaming Control

Through this chapter, we answer some of the fundamental control issues such as “how many discrete levels of engagement can a user comfortably control?” or “is it easier for users to control a parameter by becoming more engaged than less engaged?” which are largely unexplored.

Through the implemented experiments, we provide principles for one-dimensional selection in the absence of environmental or task distractions, and a final ecologically valid experiment for controlling gaming parameters through levels of engagement.

Chapter 4: Measuring task engagement as an evaluation tool for interactive tasks

In this chapter, we harness task engagement for a passive BCI-based system, where the brain activity arises without the purpose of voluntary control. We employed the continuous usage of EEG to calculate task engagement of users while interacting with an application, with D-FLIP
as an example. We investigate users’ task engagement using EEG in the context of story preparation and telling. Based on the results, we illustrate that the method for calculating task engagement can be used as an evaluation tool for interactive tasks.

**Chapter 5: Error-Related Negativity for Single User Interactive Applications**

The contributions of this chapter include describing an online single trial ERN detection technique that was verified through data acquired from the frontal-central cortex of the human brain. This method was used to detect ERN online on a single trial in an object selection task. This demonstrates the abilities of harnessing ERN in interactive applications in office conditions. Both our experiments show that we can detect ERNs using the Emotiv headset with an accuracy of up to 70%. These rates are indicative of the type of accuracy one can expect from off-the-shelf EEG sets. Through the results of the final experiment in chapter 5, these detecting rates were shown to be acceptable for interactive applications.

**Chapter 6: Error Related Negativity in Collaborative Applications**

This chapter demonstrates that observer ERN can be detected in a single trial and online basis, with the accuracy up to 64%, using a commodity headset (e.g. Emotiv EEG). In addition, the anticipation effect in observing tasks when one person observes another person committing errors was investigated using both clinical and commodity headsets. The results show that there are ERN-like patterns detected in the observer’s EEG using clinical and commodity headsets. This is a demonstration of the anticipation effect where the observer is aware of the going-to-be error of the performer.

**7.1.2. Doing it differently**

The user studies described in this thesis were implemented with some amount of the experimental trial and error. At times, the design decisions were limited by the availability of and access to the latest technologies and techniques. Therefore it is possible to recommend some changes to the design approach taken for designing and implementing the proposed BCI-based interactive applications. There are desirable steps for improving the existing systems in light of more advanced tools and technology become available. These changes do not affect the original objectives and aims of the studies. They can be treated as upgrades that could be undertaken as future work for an improved version of the investigations.


**a. Hardware upgrades**

In 10 out of 11 user studies in this thesis, we used Emotiv to capture raw EEG signals for the purpose of calculating Task Engagement or detecting ERN. This device is portable and affordable for the majority of users or gamers to harness the benefits of BCI-based interactive tasks. However, this device has fixed channel locations, which limits the research possibilities to a specific area.

An example of a direct upgrade of the hardware would be the Emotiv Insight [4], a newly design of the Emotiv headset. This headset features dry electrodes, which are more hygienic and more comfortable for users. Also, the built-in 6-asis inertial sensor and 3-axis magnetometer can be used to enrich the tasks assigned to the users and broaden the input bandwidth of the task.

Another advance on EEG devices that attracts interest from research is the development of contactless EEG electrodes (for example, [174] proposes an initial step toward a contactless ECG and EEG). Although there is not a contactless EEG system available on the market, we believe there will soon be. This will boost the interest of ordinary users toward BCI-based input due to a less complicated setup and less interruption during task performance. These changes however will require more customization in task design based on the current design guidelines.

**b. Software upgrades**

*Quantifying EEG Measured Task Engagement for use in Gaming Applications*

The software upgrades related to the algorithm to calculate task engagement and the visualizations during the experiments. For each EEG sample captured, a task engagement was generated in the [0, 1] range based on alpha, beta, and theta power for appropriate channels. All of the steps including signal capturing, signal processing, and generating task engagement values were manually coded using Visual C#. This was restricted to particular programming languages because of the available SDK at that time. However, this step can now be auto-processed using OpenVibe [155] or BCI2000 [162] which not both support Emotiv. Using these software not only helps automating the procedure but also make it easier to use a different EEG headset as long as it is supported. Currently OpenVibe and BCI2000 support the most common devices that make the procedure of using new EEG headset easier and less time intensive.

*Measuring Task Engagement as an evaluation tool for interactive tasks*

As described in chapter 4, our algorithm can be written in any language, which allows D-FLIP to be implemented in various platforms of different devices. For example, D-FLIP can be
implemented and customized for a multi-touch screen, e.g. 24 inches or larger. This enriches the story sharing experiences and may increase comfort of users as they can point and manipulate directly the photos on the screen. However, whether task engagement increases or not remains as an unanswered question due to the increase in hand movement between rest positions and the touch screen.

Similar to the previous paragraph, the process of capturing signals, signal processing, and calculating task engagement for two participants at the same time can be passed to BCI software suite such as OpenVibe or BCI2000. This helps to replicate the processing of using this evaluation method to evaluate other interactive applications.

Detecting Error Related Negativity for Interaction Design

The studies in this chapter aimed to investigate the opportunities of harnessing ERN for interactive tasks. Consequently the first step we took was to repeat the flanker task. The second step aimed to detect ERN in an interactive task in between the flanker task and a more colourful application. An alternative that was considered was to design the programming task so that the code section is bookmarked and synchronized with offline data (e.g. moments when an ERN is detected). This could then be compared with the ground truth, the correctness of the code itself in order to get the ERN detection rate. Doing this helps bringing the task online where ERN detection can be used directly in real time.

Error Related Negativity in observing interactive tasks

In this chapter, we captured EEG from the observer aiming to detect ERN when the executer of a task committed a mistake. The upgraded version would be to capture EEG signals from both observer and executer during the tasks. This could be used to analyse the combined ERN detection rates from both participants. However, the task would then need to be modified so that the executer commits mistakes (or incorrect decisions) naturally in order to trigger the ERN in the executer’s neuronal activities. This improvement would shorten the gap to harness ERN in HCI scenarios, particularly in collaborative environments around tabletops.

7.2. Discussions and future works

This thesis describes novel interaction designs using EEG for active and passive control methods. However, there are many challenges left to bring EEG signals to everyday use of HCI. Here we discuss some details and possible directions for future works.
7.2.1. Quantifying Task Engagement for Gaming Control

a. Continuous vs. Discrete

The assignment of the engagement channel to a control should also characterize the selection of continuous or discrete input. For example, weapons selection in a game would normally require discrete input control, but games with BCI could use continuous input to provide weapons that, in between discrete states, act as a morphed augmentation of two weapons. Where BCI is used as a passive parameter manipulator, the engaged or disengaged brain states could adjust the balance between armour or weapon strength; this way, users who struggle with exact parameter selection do not ‘lose out’ in a BCI augmented game.

It is also possible to design a game with characters adapting to different levels of engagement. For example, in a game like Warcraft Dota Allstars (www.playdota.com) where characters have intelligence, strength and agility characteristics, and low engagement could increase strength; medium engagement could increase intelligence and high engagement could increase the character’s agility.

b. Serial vs. Parallel Interaction

There are no technical restrictions on continuously monitoring and interpreting EEG signals. However, unlike a mouse or keyboard it is unclear when input starts and ends. EEG input is ‘always on’, so designers must consider when to utilize this input stream. For BCI-assisted games, there are two options: serial or parallel interaction.

Serial interaction allows gamers time to concentrate on producing the engagement level required to complete the task - they do not need to perform mouse or keyboard input at the same time. Once their engagement level is measured they continue playing the game using more traditional input methods. Serial interaction may allow users to complete their tasks faster (as it has their full attention), but breaks the game’s flow. For example, in a game like Delta Force a sniper’s target aim can be proportional to their level of engagement. If a player needs to aim at a small or far target he would need to bring down his level of engagement to the lowest level before firing the gun.

In contrast, parallel interaction requires the user to continue input using traditional controllers, while appropriately adjusting their engagement level to match the level required by the game. Parallel input increases the number of simultaneous inputs and allows a smoother flow of gameplay. However, performing other tasks using traditional inputs may distract gamers who are
concentrating on controlling their level of engagement. Consequently, users may take longer or struggle to complete BCI-controlled tasks. For example, in a game like Diablo the player can move on a map and cast a spell simultaneously. The strength of the spell can be controlled by the engagement level and if the character is too close to the monster then a strong spell can also kill the player. This would encourage diligence in selecting the right spell level (encouraging control of engagement) but would also support error resilience by allowing players to kill a monster through multiple lower-level spells.

c. Generalizability of Results

We achieved our initial goals of providing more general guidelines for interaction with EEG signals, as demonstrated through the various possibilities for game design, for example. We specifically considered the sensed parameter of engagement. We used Emotiv’s SDK to access EEG signals in real-time to carry out our experimental studies. Our studies were based around the raw EEG signals to provide the most flexible set of guidelines that can then be applied to other EEG toolkits such as MindWave from Neurosky™, which also provide access to raw signals. One limitation is the need to calibrate the EEG signal to each user’s engagement ability. While this may be somewhat mechanical for our studies, we can better integrate this through an initial game-play. This would need to be done only once and can be changed adaptively over time.

7.2.2. Measuring task engagement for interactive task evaluation

As demonstrated in chapter 4 with D-FLIP as an example of an interactive application, task engagement can be measured passively in order to understand users’ underlying states with the purpose of enriching interactive experiences. We suggest some potential applications that can benefit from our findings as follows.

a. Enhance interactive experiences with eye tracking devices

The results from our experiments showed that users could comfortably interact with the application while their EEG signals are being captured, using a commodity EEG headset. This type of headset can also be used with other tracking devices to have more insights into users’ behaviours and cognitive states. For example, Emotiv can be combined with an eye tracking device to synchronize the information of where the users were looking and their task engagement levels around those moments. This is useful in several contexts such as objects / tasks organizing on screen when users multi-task. One of the tasks may require much more task engagement than other tasks due to the large amount of workload. The system provide the users
with suggestion such as automation some tasks or adjusting the visualization so that the users can focus on a smaller number of tasks or on a single well-adjusted task in order to reduce the workload. On the other hand, a very low detected task engagement means that the ongoing task makes the user feel bored while interacting. The system can then increase the workload or ask the user to manage more tasks, in a way that increases task engagement without increasing the user’s anxiety.

b. Generalization of task engagement values
As pointed out earlier, a good interactive application should keep users’ engagement level not too high to make users anxious but not too low, as the users can feel bored. However, exact percentages or numbers on task engagement range have not been investigated. Therefore, a further research on building a database of task engagement ranges according to user types and task types could potentially solve this problem. Using the database, we could determine what the expected optimal task engagement values when a user is interacting with the application.

7.2.3. Detecting Error-Related Negativity for Interaction Design
As ERN is elicited in any multiple choice RT time task when the user is confused or aware of an accidental action they have made. This has great potential in many types of interactive applications. We suggest some of them in the following sections.

a. Gaming conditions
ERN can be used to provide users with a new form of experience in gaming. For example, in time critical missions (like shooting a character in an online game) sometimes network delays and other external factors may affect the overall outcome of the mission or battle. In these cases if an ERN is detected this can be used in systems decision making process either by giving the player another chance or changing the time-stamp of user triggered events to make the outcome seem as intended by the user.

b. Aiding Object Selection
Selecting a static target among a selection of objects is a multiple choice reaction time task. The user starts with deciding on the target then makes an initial open-loop movement followed by a final correction phase where they move the pointer or finger on the target to select it. In this correction phase the user usually receives visual feedback on whether or not their selection was successful. This feedback combined with the initiation of the correction-phase movement can trigger ERNs if the movements are fast enough. Thus it should be possible for an ERN detection
module to detect errors in users' intention and attempt to correct it. In many instances the cost of recovering from a wrong button press or a pointer selection can be quite high – the application might be launched and the user would have to close it before re-launching the right application. In these circumstances when ERN is detected the system can prompt the user if the target was selected correctly potentially helping the user.

ERN detection can be combined with P300 to reduce the user’s mental load and frustration (error-recovery is time consuming and difficult) that is associated with working with P300 data. For example, in an object selection using P300 on a multi-touch table [197], ERN can be used as final confirmation that the object is the user's intended selected object. This will be very useful because it is difficult to confirm using only P300 that the selected object on the table is the one that the user wanted. Moreover, ERN usage can eliminate the trial of hitting the BACK button to de-select an object in P300 spelling which is time consuming and requires high concentration. Additionally, this can also be applied into object selections on a tabletop for people in working condition so that the time spent in correcting wrong selection will be reduced and efficiency improved.

c. Mobile Spatial Navigation

Sometimes, a user still needs help in using and navigating using electronic maps on the fly (i.e. Google maps). HCI designers may integrate ERN into the system so that it can detect the confusion and error awareness moments in order to provide appropriate suggestion based on the location. One obstacle is EEG is known to be sensitive to movements such as walking and moving your body. This can make the EEG signal very noisy reducing the accuracy of the ERN detector. Before being fed into the classifier, EEG signals may need to be carefully pre-processed.

d. Multiple users' applications

ERN use can be extended to multiple users scenarios where a person's ERN is made visible to the entire team so team-support is available when the user is confused. For example if a gamer is confused in navigation or shooting activity, his/her teammates can assist him, or in collaborative tabletop applications, other people may give help and suggestion to a person whose avatar is being shown as confused or aware that they have made a wrong selection encouraging collaborative peer-learning.
7.2.4. Error-Related Negativity in Observing Interactive Tasks

Besides the benefit of correcting one’s own errors, ERN has the potential to enrich interactive applications in the collaborative working environment. The results from the studies in this chapter can provide guidance on how best to begin harnessing ERN for such interactive experiences. From our results, HCI designers can employ ERN in interactive tasks to pinpoint the whereabouts of an error, caused by the performing user, which can be from 50ms before to 250ms after the action that triggered it, depending on the designed interaction technique.

The usefulness of executer’ and observer’ ERNs in interactive tasks depends on the usage context and the designer’s creativity in making use of it. Here we suggest some applications to highlight design possibilities.

a. Collaborative tabletop settings

Groupware applications use embodiments to help people stay aware of the presence, location and movement of users in collaborative tasks [82]. The embodiments can be real embodiments which make use of the actual physical body of the user [84] or virtual embodiments which are digital representations of users (e.g. Telepointers [81]).

In these settings, both real and virtual embodiments provide obvious awareness information to other users in the group. As a result, an observer can judge quickly about the correctness of the action being performed. Moreover, users usually aim to touch the destination point as quickly as possible once the decision is made in their minds. This fits perfectly as a multiple-choice RT task that has high chance of eliciting an ERN pattern in the observers’ brain. However, this information is not accessible unless the observers provide feedback to either the system or the performer. This process is time consuming and can interrupt the performance.

Our results show that ERN detection module can be integrated to monitor the appearance of ERN patterns in the observers. As the classification time is about 1ms with the coefficient matrix known in advance, the detected ERN can be used to trace back the time period where the observer thinks the performing user made a mistake or is confused about their action. Although these time periods range from 50ms before to 250ms after the committed action, careful task design can help to shorten this range to specific time points. This is due to different interaction techniques providing different awareness to the observers leading to different ERN triggered moments [135]. Hence, HCI designers should design their applications so that action awareness is maximized to trigger clearer ERNs.
As a simple example, a performer’s action that leads to an observer ERN can result in an ignorable pop-up menu that can help revert or discuss based on the source of ERN in the observer. Rather than merely automate the reversion of an action, ERN can serve as an information point for discussion and clarification between the collaborators. The pop-up menu can then provide contextually relevant information including an “undo” item.

The goal is to maximize group awareness of individual’s actions without disrupting the workflow of the task. Further contextual studies can explore this trade-off.

b. War-rooms / Emergency response

Giving instructions and making decisions is crucial in emergency scenarios. ERN can be integrated into the Emergency Management Information System [107] so that each decision that has been made can be cross checked by the supervisors and observers in the same team. If the executer is made aware that other teammates think a mistake is about to be made by the performing action, it can be prevented before it is made. Moreover, if the mistake is already made, the system can detect if an ERN appears in an observer’s brain to speed up the correction.

In emergency response settings where time and accuracy is key, a clinical EEG headset might be useful. On the other hand, using commodity headsets such as Emotiv does provide a range of benefits and flexibility that might make it more useful. For example, clinical headsets are cumbersome and limit mobility of the observers while a commodity headset although slightly slower in capturing ERN is mobile and allows the observer to dynamically respond to an emerging situation. With the fast detection time of the ERN detection module, short detection time can be used as a buffer before an action is executed potentially preventing serious errors. A consequence of such designs that needs careful consideration is the training offered by such a system to the performer. Without careful design, the system can encourage the performer to absolve them of responsible action (passing responsibility to the observer) however good design can lead to better training of the performer (as explored in competitive gaming environments).

c. Gaming scenarios

Usually in MMORPGs (e.g. World of Warcraft) and shooting games (such as CounterStrike or Half-Life) gamers team up to act against other teams. During this period, every decision needs to be precise and made in a timely manner. ERN could be integrated to give team members a new tool to observe and react to each other’s actions. Combinations of detected ERNs from teammates can be displayed as feedback to every continuous action say in the form of an overlay window. This can give gamers an idea of where and when they committed a mistake,
which could add a new dimension to the game tactics. This ERN communicating channel, beside verbal and chatting communication, can be used to improve the teamwork and strategic analysis skills during the game hence make it more challenging between teams.

d. Hyperscanning

In collaborative scenarios, the system can check for ERNs in multiple observers to improve the robustness of the detection module. As the ERN patterns appear with similar latency for each task settings, EEG signals can be measured from multiple people at the same time to do EEG hyperscanning [19] where signals from multiple users can be fed to a single classifier. As a result, this can improve the classification rate because of more data to train the classifier and obtain a clearer ERN. It leads to an increase of the robustness and classification rates of the ERN detection module. This method can also speed up the detection process by using a shorter input window size hence reduces the classifier limitation noted above.

7.2.5. Vision of the future works

The works in this thesis focus on investigating capabilities of designing interactive applications using active and passive BCI-based systems. We proposed several guidelines for HCI designers to exploit our results. However, here we only examined specific neural states (such as ERN) with many other states remaining open for further investigation (such as imaginary movements and N100). We envision future researchers will fill this gap. Furthermore, with the technologies advancing forward quickly, EEG devices will become more pervasive (readily available everywhere) and less invasive (through advances in contactless EEG sensing). This new development will open many more applications that can benefit from both passive and active BCI-based systems leading to research about increased input bandwidth as well as their associated design guidelines. In addition to that, with more and more research into signal processing and pattern recognition, robust and noise-tolerated sensing will become readily available. In this case, our rather conservative approach to proposing design guidelines can be extended with guidelines that are a consequence of lower detection errors. Consequently, the overall aim is to make BCI systems be accessible and enjoyable to majority of users with this new type of intuitive input.

7.3. Conclusions

Through several investigations in this thesis, we have demonstrated how to calculate task engagement for the purpose of an additional control input for HCI as well as to evaluate interactive applications. We also presented how to detect ERN patterns on an online, single trial
basis which with the detection accuracy that is beneficial to HCI. These are examples of employing EEG for active and passive BCI systems. In addition, we showed that it is possible to design EEG-based systems using off-the-shelf EEG headsets, such as Emotiv.

We also analysed design guidelines on how our results can be harnessed to benefit interactive tasks through detailed discussions. In addition, we explored many directions for future work in terms of evaluating interactive tasks, detecting awareness of errors and their correction in both single- and multi-user scenarios. This thesis is one of the initial attempts to bring BCI closer to the majority of users and closer to Human Computer Interaction.
Appendices

1. Emotiv EEG Neuroheadset Specifications

<table>
<thead>
<tr>
<th><strong>EMOTIV EEG HEADSET</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of channels</td>
<td>14 (plus CMS/DRL references, P3/P4 locations)</td>
</tr>
<tr>
<td>Channel names (International 10-20 locations)</td>
<td>AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4</td>
</tr>
<tr>
<td>Sampling method</td>
<td>Sequential sampling, Single ADC</td>
</tr>
<tr>
<td>Sampling rate</td>
<td>128 SPS (2048 Hz internal)</td>
</tr>
<tr>
<td>Resolution</td>
<td>14 bits 1 LSB = 0.51µV (16 bit ADC, 2 bits instrumental noise floor discarded)</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>0.2 - 45Hz, digital notch filters at 50Hz and 60Hz</td>
</tr>
<tr>
<td>Filtering</td>
<td>Built in digital 5th order Sinc filter</td>
</tr>
<tr>
<td>Dynamic range (input referred)</td>
<td>8400µV (pp)</td>
</tr>
<tr>
<td>Coupling mode</td>
<td>AC coupled</td>
</tr>
<tr>
<td>Connectivity</td>
<td>Proprietary wireless, 2.4GHz band</td>
</tr>
<tr>
<td>Power</td>
<td>LiPoly</td>
</tr>
<tr>
<td>Battery life (typical)</td>
<td>12 hours</td>
</tr>
<tr>
<td>Impedance Measurement</td>
<td>Real-time contact quality using patented system</td>
</tr>
<tr>
<td>Price</td>
<td>299 ~ 750 USD</td>
</tr>
</tbody>
</table>
## 2. EbNeuro BE MRI System Specifications

<table>
<thead>
<tr>
<th></th>
<th>EbNeuro BE MRI System</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of channels</strong></td>
<td>21</td>
</tr>
<tr>
<td><strong>Used channels</strong></td>
<td>Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2</td>
</tr>
<tr>
<td><strong>Sampling rate</strong></td>
<td>Programmable (up to 8 kHz/ channel) and multi-frequency</td>
</tr>
<tr>
<td><strong>Resolution</strong></td>
<td>14 bits 1 LSB = 0.51µV (16 bit ADC, 2 bits instrumental noise floor discarded)</td>
</tr>
<tr>
<td><strong>Bandwidth</strong></td>
<td>DC-20 kHz</td>
</tr>
<tr>
<td><strong>Filtering</strong></td>
<td>Built in digital 5th order Sinc filter</td>
</tr>
<tr>
<td><strong>Dynamic range (input referred)</strong></td>
<td>128 mV</td>
</tr>
<tr>
<td><strong>Power</strong></td>
<td>Main power</td>
</tr>
<tr>
<td><strong>Impedance Measurement</strong></td>
<td>Ohmmeter Threshold Led to check electrode montage impedance directly on the head-box</td>
</tr>
<tr>
<td><strong>Signal Quality</strong></td>
<td>High (low noise, high sensitivity for a large dynamic, high CMRR)</td>
</tr>
<tr>
<td><strong>Artefact</strong></td>
<td>Artefact rejection, patient safety and versatility(connecting solutions allow to place the acquisition unit 500 m. far from the controlling workstation)</td>
</tr>
<tr>
<td><strong>Burst suppression</strong></td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Price</strong></td>
<td>14,041.41 USD</td>
</tr>
</tbody>
</table>
3. Experimental Consent Form

Department of Computer Science
Interaction & Graphics Group

CONSENT FORM

Please answer the following questions to the best of your knowledge

<table>
<thead>
<tr>
<th>HAVE YOU:</th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>• been given information explaining about the study?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• had an opportunity to ask questions and discuss this study?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• received satisfactory answers to all questions you asked?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• received enough information about the study for you to make a decision</td>
<td></td>
<td></td>
</tr>
<tr>
<td>about your participation?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• ascertained that you don’t have any known condition that prevents you</td>
<td></td>
<td></td>
</tr>
<tr>
<td>from taking part in this study?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| DO YOU UNDERSTAND:                                                                 |
|-------------------------------------------------------------------|-----|----|
| that you are free to withdraw from the study and free to withdraw  |     |    |
| your data prior to final consent                                  |     |    |
| • at any time?                                                    |     |    |
| • without having to give a reason for withdrawing?                |     |    |

I hereby fully and freely consent to my participation in this study

I understand the nature and purpose of the procedures involved in this study. These have been communicated to me on the information sheet accompanying this form.

I understand and acknowledge that the investigation is designed to promote scientific knowledge and that the University of Bristol will use the data I provide for no purpose other than research.

I understand the data I provide will be **anonymous**. No link will be made between my name or other identifying information and my study data.

I understand that the University of Bristol may use the data collected for this study in a future research project but that the conditions on this form under which I have provided the data will still apply.

I agree to the University of Bristol keeping and processing the data I have provided during the course of this study. I understand that these data will be used only for the purpose(s) set out in the information sheet, and my consent is conditional upon the University complying with its duties and obligations under the Data Protection Act.

<table>
<thead>
<tr>
<th>Participant’s signature:</th>
<th>Date:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Name in BLOCK letter:
4. NASA-TLX Questionnaire

Hart and Staveland's NASA Task Load Index (TLX) method assesses workload on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

<table>
<thead>
<tr>
<th>Name</th>
<th>Task</th>
<th>Date</th>
</tr>
</thead>
</table>

**Mental Demand**
How mentally demanding was the task?

| Very Low | Very High |

**Physical Demand**
How physically demanding was the task?

| Very Low | Very High |

**Temporal Demand**
How hurried or rushed was the pace of the task?

| Very Low | Very High |

**Performance**
How successful were you in accomplishing what you were asked to do?

| Perfect | Failure |

**Effort**
How hard did you have to work to accomplish your level of performance?

| Very Low | Very High |

**Frustration**
How insecure, discouraged, irritated, stressed, and annoyed were you?

| Very Low | Very High |
5. Information sheet for D-FLIP

Research Institute of Electrical Communication
2-1-1 Katahira, Sendai, Miyagi Prefecture

Healthy volunteer information sheet – Dynamic PhotoViewer study
You are being invited to take part in a research project. Here is some information to help you decide whether or not to take part. Ask us if there is anything you do not understand or if you would like more information on any aspect of the project. Take time to decide whether or not you wish to take part.

What is the purpose of the study?
In this study, we aim to investigate the differences in users’ relaxation and engagement while using our new application (PhotoViewer) and a standard program (Windows Explorer). Users’ relaxation and engagement will be measured in real-time using EEG. The scenario will be in a storytelling context.

Do I have to take part?
No, taking part is voluntary. It is up to you to decide whether or not to take part. If you do decide to take part we will ask you to sign a consent form and give you a copy of this information sheet and the consent form to keep. If you decide to take part you are still free to withdraw at any time. If you decide not to take part you do not have to give a reason, nobody will be upset.

The study will be carried out at Room S20, Kitamura lab, Research Institute of Electrical Communication, 2-1-1 Katahira, Sendai, Miyagi Prefecture.

What will happen to me?
We have given you this information sheet since we have found that you are potentially suitable to take part and this sheet gives you additional information on the study and will help you to decide if this is something you are interested in doing. If you decide to take part, you will be invited to discuss the study further and provide written consent.

During the experiment, you will wear an EEG cap which is pain-free and risk-free for your head, hair, or skin (see Figure 1). Some solution is placed under the cap at some locations. This solution is approved to be used in the study and easily wash away.
What are the possible disadvantages and risks of taking part?
There is no risk for the participants during the experiment. All equipment is approved to be used in clinical studies. The device only read EEG signals. It does not put any signals back inside your brain.

Confidentiality – Who will know I will be taking part?
All information collected about you during this study will be kept strictly confidential.

What will happen to any data that I give?
The collected data will be stored for analysis, up to 60 months. After that period, data will be destroyed.
If you decide to withdraw from the study, you can ask for your samples to be destroyed if they have not already been analysed, or for any information obtained from analysing your data to be destroyed.

What will happen to the results of the study?
Results will be submitted to scientific journals for publication, and presented at scientific conferences.

What do I need to do during the experiment?
During the experiment, you will be required to wear an EEG headset for the duration of the experiment. This headset does not cause any damage to your head, brain, or hair. In addition:

If you are the teller participant:
- Before coming to the experiment, you will be asked to bring 2 sets of personal photos which were taken in your previous trips. Each set should contain at least 200 photos. They will be resized to a maximum size of 640 x 640 pixels.
- **Step 1:** During the experiment, you will be asked to select 10 photos of each set and form a story from them. The minimum time to spend in this step is 5 minutes. You are asked to use each of two applications to do this task (see Figure 2 and Figure 3).
- **Step 2:** After that, you will tell your story to another person. You can still use the application to enrich your story.

If you are the listener participant:
- You will watch how another participant select their photos in step 1. And
Figure 2: 2-Windows Story Telling Mode

Figure 3: Dynamic PhotoViewer story telling mode

Please remember that during the experiment:

- Try to minimize your body, head movements, and facial actions (smile, laugh,…) during the experiment
- For the teller participant: you can move freely in step 2 (when telling the story).

What do I do now?

If, after reading this information, you decide that you would like to take part, please contact one of the study researchers and we will arrange some convenient times for you to attend.

We will be happy to answer any questions you have. If you are prepared to take part, you will be asked to sign a consent form to confirm this. You will be given this information sheet to
keep. We suggest you keep it carefully so that you can contact us if you have any further questions, at any time.

You do not have to take part, and if, at any stage during the study you wish to withdraw, for whatever reason, you are able to do so.

**Contact for further information**

For any further information on this study you can contact one of the study researchers:

Dr. Kazuki Takashima Tel: 022-217-6119 takashima@riec.tohoku.ac.jp

Dr. Hitomi Yokoyama Tel: 022-217-6119 yokoyama@riec.tohoku.ac.jp

**What if there is a problem?**

If you have a concern about any aspect of this study, you should ask to speak with the researchers who will do their best to answer your questions: Chi Vi (vi@cs.bris.ac.uk) or Dr. Kazuki Takashima (takashima@riec.tohoku.ac.jp).

If you wish to make a formal complaint, please write to:

Research Institute of Electrical Communication, 2-1-1 Katahira, Sendai, Miyagi Prefecture.

Thank you very much for considering taking part in our research. Please discuss this information with your family and friends if you wish.
6. Information sheet for Task Engagement experiments

In this study, you will be asked to wear an Emotiv Headset (Figure 1). This device is non-invasive (measure EEG – brain signals by attaching electrodes to the skin).

Note: During the experiments, if you feel uncomfortable or do not want to continue, just stop and tell the experimenter.

Figure 1. Emotiv Headset

- It is not harmful to your head.
- It just read the signals, not put anything into your head!
- It does NOT read your thoughts!

The purpose of this study is to investigate how well you can control your own engagement levels. It is divided into three steps:

- **Step 1:** Measure EEG signals when you close your eyes and relax.
- **Step 2:** Measure EEG signals when you do a task requires high engagement.
- **Step 3:** Use your engagement to reach a certain level.

Note: During the experiment, please minimize body movements, head movements and facial actions (laughing, clenching, talking) unless it's critical.

Step 1:

Figure 2: Eye close – Low engagement

- While wearing the Emotiv headset, sit comfortably, and close your eyes will the experimenter tell you to do so (around 5 minutes).
- Try to clear your mind and relax
- Do not fall asleep!!!
- Open your eyes when the experiment tells you to do so.
Step 2:

![Three Choice Vigilance Task](image)

**Figure 1. Three Choice Vigilance Task**

The procedure of this task is:

- One of three objects:
  - 1 Primary (RED)
  - 2 Secondary (Yellow or Green)

will appear on the screen for 0.2 seconds. Your task is to press the SPACEBAR by both index fingers AS SOON AS POSSIBLE after you saw the primary object ONLY (the red rectangle).

- If you do not press SPACEBAR within 2 seconds after a primary object, the trial will be marked as TIMEOUT. Otherwise, it will be marked as COMPLETED.
- Do not press SPACEBAR after a secondary object.
- After each trial, there will be a few seconds waiting before a new trial starts.
- This experiment lasts (around 5 minutes).

Step 3:

- Your task is to move the car from start point (on the left) to the Finish point (on the right).
- A trial starts by asking you to adjust your engagement level to a certain level (either level 0 or level 4 varies between each trials). When your current engagement level and the start level overlap for 0.1s. You then have to move the car from start point to finish point in 30s.
- The car will move forward when you press the RIGHT ARROW key on the keyboard.
- When the car moves ¼ of the way, a monster appears and fires a fireball toward the car (the red ball) with a number inside.
- You need to adjust your engagement level (yellow colour) to match the number inside that fireball. When those two level match, theirs colour will change to Red. At that moment, press F5 key on the remote to confirm.

Note:

- If you cannot reach Finish point in time, the trial is failed.
- If you reach the Finish point but failed to destroy the fireball, the trial is failed.
- You can press the F5 key multiple times, even before or after they match. However, that key press is only recorded when those two colours match.
- The time to move from start point to finish point is about 20s.
- The fireball will hit the car in 10s if you cannot adjust your engagement level to that level.
- It is recommend to adjust your engagement level to destroy the fireball while continue moving the car to save time. However, you can stop and destroy the fireball, then move forward to the Finish point.
7. Information Sheet for Flanker Task

In this experiment, we are investigating the error awareness in an observing task. You will be asked to perform a version of Flanker task. After that you will observe another person performing 4 blocks of that task while giving feedback about their performance of each trial.

You will be required to wear an Emotiv headset for the duration of the experiment. This headset does not cause any damage to your head, brain, or hair.

**Step 1**

![Figure 1: Flanker stimulus](image)

Each trial starts by displaying a Flanker stimulus consisted of 5 arrows. This stimulus will appear at the centre of the screen for 0.3s, and then disappear.

You need to respond based on the pointing direction of the middle arrow by pressing LEFT or RIGHT button on a remote. For example, press RIGHT button for the stimulus in Figure 1.

![Figure 2: Response](image)

Your answer will be display on the screen for 1s (Figure 2).

After that, you will be asked how confident you are with the last decision (Figure 3).

![Figure 3: Question about the last decision](image)

Choose your answer by pressing the appropriate key on the keyboard.

**Step 2**

In this step, you will observe another person performing the task. After each trial, you will answer the question (Figure 3) about the correctness of the last action by press the appropriate key on the keyboard.

*Note: During the experiment:*
- Try to minimize your body, head movements, and facial actions (smile, laugh, ...) during the
- Try not to blink right before, during and right after the Response (Figure 2) is displayed.
8. Information Sheet for ERN in Observation Task

Healthy volunteer information sheet – BCI tabletop study

You are being invited to take part in a research project. Here is some information to help you decide whether or not to take part. Ask us if there is anything you do not understand or if you would like more information on any aspect of the project. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

In this study, we aim to investigate the moments of error awareness and confusion following a fast response in an observation task. Previous study have shown that when an person observe another person making an incorrect choice in a reaction time task, there is an unique pattern named Error Related Negativity appear in the EEG (electrical signals emit by neurons). We aim to investigate further in the scenario of interacting around an interactive tabletop.

Do I have to take part?

No, taking part is voluntary. It is up to you to decide whether or not to take part. If you do decide to take part we will ask you to sign a consent form and give you a copy of this information sheet and the consent form to keep. If you decide to take part you are still free to withdraw at any time. If you decide not to take part you do not have to give a reason, nobody will be upset.

The study will be carried out at University of Bristol, Clinical Research and Imaging Centre (CRICBristol), 60 St Michael's Hill, Bristol, BS2 8DX

What will happen to me?

We have given you this information sheet since we have found that you are potentially suitable to take part and this sheet gives you additional information on the study and will help you to decide if this is something you are interested in doing. If you decide to take part, you will be invited to discuss the study further and provide written consent.

During the experiment, you will wear an EEG cap which is pain-free and risk-free for your head, hair, or skin (see Figure 1 for example). Some clinical EEG gels are placed under the cap at some locations. These gels are approved to be used in the study and easily wash away. The cap we use in this study is similar to the one in Figure 1.
What are the possible disadvantages and risks of taking part?

There is no risk for the participants during the experiment. All equipment is approved to be used in clinical studies. The device only read EEG signals. It does not put any signals back inside your brain.

There might be a little discomfort (and a cool feeling) when we put the cap on as the gels touch your skin for the first time.

**Will I receive any reimbursement for my time?**

Yes. You will receive £10 as our gratitude for your contribution.

**Confidentiality – Who will know I will be taking part?**

All information collected about you during this study will be kept strictly confidential.

**What will happen to any data that I give?**

The collected data will be stored for analysis, up to 60 months. After that period, data will be destroyed.

If you decide to withdraw from the study, you can ask for your samples to be destroyed if they have not already been analysed, or for any information obtained from analysing your data to be destroyed.

**Who is organising and funding the research?**

The Department of Computer Science which is part of the University of Bristol is carrying out the research. The research is currently funded by resources within the University. Funding pays the salaries of research staff and other direct costs of doing the research. Researchers are not receiving any payments other than their usual salaries.

**What will happen to the results of the study?**

Results will be submitted to scientific journals for publication, and presented at scientific conferences.
What do I need to do during the experiment

You will be asked to perform a version of Flanker task. After that you will observe another person performing 4 blocks of that task while giving feedback about their performance of each trial.

You will be required to wear an EEG headset for the duration of the experiment. This headset does not cause any damage to your head, brain, or hair.

Step 1

Figure 2: Flanker stimulus

Each trial starts by displaying a Flanker stimulus consisted of 5 arrows. This stimulus will appear at the centre of the screen for 0.3s, and then disappear.

You need to respond based on the pointing direction of the middle arrow by pressing LEFT or RIGHT button on a remote. For example, press RIGHT button for the stimulus in Figure 2.

Figure 3: Response

Your answer will be display on the screen for 1s (Figure 3).

After that, you will be asked how confident you are with the last decision (Figure 4).

Figure 4: Question about the last decision

Choose your answer by pressing the appropriate key on the keyboard.

Step 2

In this step, you will observe another person performing the task. After each trial, you will answer the question (Figure 4) about the correctness of the last action by press the appropriate key on the keyboard. Try to remember how many times that another person did incorrectly.

Note: During the experiment:
- Try to minimize your body, head movements, and facial actions (smile, laugh,...) during the experiment
- Try not to blink right before, during and right after the Response (Figure 3) is displayed.
- Only answer the question AFTER it appears.

What do I do now?

If, after reading this information, you decide that you would like to take part, please contact one of the study researchers and we will arrange some convenient times for you to attend.

We will be happy to answer any questions you have. If you are prepared to take part, you will be asked to sign a consent form to confirm this. You will be given this information sheet to keep. We suggest you keep it carefully so that you can contact us if you have any further questions, at any time.

You do not have to take part, and if, at any stage during the study you wish to withdraw, for whatever reason, you are able to do so.

Contact for further information

For any further information on this study you can contact one of the study researchers:

Prof. Sriram Subramanian       Tel: +44 117 3315235
Dr. David Coyle                Tel: +44 117 9545468
Chi Thanh Vi                   Tel: +44 117 9545289

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak with the researchers who will do their best to answer your questions (vi@cs.bris.ac.uk) or Prof. Sriram Subramanian (sriram@cs.bris.ac.uk).

If you wish to make a formal complaint, please write to:

Research and Development

University of Bristol

Senate House, Tyndall Ave

Bristol. BS8 1TH

Thank you very much for considering taking part in our research. Please discuss this information with your family and friends if you wish.
References


98. Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E. and Sejnowski, T. J. Removal of eye activity artifacts from visual event-related


