THE USABILITY OF NATURAL USER INTERFACES FOR GAMEPLAY

by

Rouxan Colin Fouché

2002025357

Submitted in fulfilment of the requirements for the degree

Magister Scientiae

In the Faculty of Natural and Agricultural Sciences

Department of Computer Science and Informatics

University of the Free State

Bloemfontein

South Africa

2015

Study Leader: Dr T.R. Beelders

Co-Study Leader: Dr L. De Wet

Acknowledgement

I would like to express my sincere gratitude to my study leader, Dr T.R. Beelders, for the continuous support of my study and research, for her patience, motivation and enthusiasm. I would also like to express my gratitude to my co-study leader, Dr L. De Wet, for the useful comments and remarks through the learning process of this dissertation. Furthermore, I would like to thank my loved ones who have supported me through the entire process with patience and love.

Table of Contents

LIST OF TABLES VI
LIST OF FIGURES VIII
LIST OF CHARTS IX
CHAPTER 1 - INTRODUCTION 1
1.1. INTRODUCTION 1
1.2. AIM 1
1.3. PROBLEM STATEMENT 1
1.4. MOTIVATION 2
1.5. RESEARCH QUESTION 3
1.5. HYPOTHESES 4
1.6. METHODOLOGY 4
1.7. SCOPE OF THE STUDY 5
1.8. CONTRIBUTION 5
1.9. LIMITATIONS 6
1.10. OUTLINE OF THE DISSERTATION 7
1.11. SUMMARY 8
CHAPTER 2 – LITERATURE REVIEW 9
2.1. HUMAN-COMPUTER INTERACTION 9
2.1.1. INTERFACING 10
2.1.2. MODES OF INPUT 11
2.1.3. TRADITIONAL INTERFACES 15
2.1.4. NATURAL USER INTERFACES 16
2.2. BRAIN-COMPUTER INTERFACE (BCI) 18
2.2.1. DIRECT BRAIN-COMPUTER INTERACTION 18
2.2.2. BRAIN SIGNALS 19
2.2.2.1. Electroencephalogram (EEG) 19
2.2.2.2. Electrooculography (EOG) 20
2.2.2.3. Electromyography (EMG) 20
2.2.3. BCI TECHNOLOGY 21
2.2.4. BCI CLASSIFICATIONS 21
2.2.4.1. Non-Invasive Method 21
2.2.4.2. Invasive and partially invasive method 22
2.2.5. NON-INVASIVE COMMERCIALLY AVAILABLE BCIS 23
2.2.5.1. Emotiv BCI Headset 23
2.2.5.2. Neurosky 24

2.2.6. EMOTIV BCI FUNCTIONALITY 24
2.2.7. BCI APPLICATION AND FUNCTIONALITY 26
2.2.7.1. Introduction 26
2.2.7.2. BCI use for disabled and able-bodied users 27
2.2.7.3. BCI application in gaming 28
2.2.7.4. Cognitive input recognition 29
2.2.7.5. Facial expression recognition 32
2.2.7.6. Alternative cursor control 35
2.2.8. SUMMARY OF BCI TECHNOLOGY 37

2.3. GLOVE-BASED TECHNOLOGIES 38

2.3.1. INTRODUCTION 38
2.3.2. HAND POSTURES AND GESTURES 38
2.3.3. GLOVE-BASED DEVICES 39
2.3.4 APPLICATIONS FOR GLOVE-BASED DEVICES 40
2.3.4.1. Designing 40
2.3.4.2. Information visualisation 40
2.3.4.3. Robotics 41
2.3.4.4. Sign language 41
2.3.4.5. Medical applications 42
2.3.4.6. Entertainment 42
2.3.4.7. Wearable and portable computers 42
2.3.4.8. Multimodal interaction 43
2.3.4.9. Gaming applications 43
2.3.5. COMMERCIALLY AVAILABLE GLOVE TECHNOLOGY 44
2.3.6. SUMMARY OF GLOVE-BASED INPUT 48

2.4. CONCLUSION 48

CHAPTER 3 - RESEARCH DESIGN 50

3.1. INTRODUCTION 50

3.2. RESEARCH PROBLEM 50

3.3. RESEARCH QUESTION 50

3.4. EXPERIMENTAL DESIGN 50

3.5. USABILITY TESTING 52

3.5.1. USABILITY 54

3.6. SURVEYS 57

3.6.1. QUESTIONNAIRES 58
3.6.2. INTERVIEWS 59

3.7. SAMPLING 59

3.6.1. NONPROBABILITY SAMPLING 60
3.6.2. PROBABILITY SAMPLING 61

3.8. CONCLUSION 61

CHAPTER 4 - METHODOLOGY 63

4.1. INTRODUCTION 63

4.2. EXPERIMENTAL SETUP 63
4.3. INTERFACE COMBINATIONS 65
4.4. USABILITY METRICS 68
4.5. PILOT STUDY 70
4.6. USER TESTING 70
4.6.1. POPULATION, SAMPLING AND SAMPLE SIZE 70
4.6.2. TEST SESSIONS 72
4.7. ISSUES OF RELIABILITY AND VALIDITY 73
4.8. DATA ANALYSIS AND INTERPRETATION 73
4.9. CONFIDENTIALITY AND ETHICAL ISSUES 74
4.10. CONCLUSION 74
CHAPTER 5 – PILOT STUDY 75
5.1. INTRODUCTION 75
5.2. PILOT STUDY ANALYSIS 75
5.2.1. TOTAL TIME 76
5.2.2. INITIAL CONTACT TIME 77
5.2.3. TIME ON TARGET 78
5.2.4. NUMBER OF OVERSHOOTS 79
5.2.5. ANECDOTAL OBSERVATIONS 81
5.3. CONCLUSION 81
CHAPTER 6 - ANALYSIS 83
6.1. INTRODUCTION 83
6.2. ANALYSIS 83
6.2.1. REPEATED MEASURES ANOVA 83
6.2.2. ANALYSIS OF SEPARATE TASKS 84
6.2.3. INTERACTION TECHNIQUES 85
6.3. TOTAL TIME TO COMPLETE THE TASK 85
6.4. RELOAD TIME 91
6.5. INITIAL CONTACT WITH TARGET 95
6.6. NUMBER OF OVERSHOOTS 105
6.7. TIME ON TARGET 116
6.8. NUMBER OF MISSES 124
6.9. NON-ELIMINATED TARGETS 130
6.10. RESETS 131
6.11. ANALYSIS OF SUBJECTIVE MEASURES 135
6.11.1. SATISFACTION 135
6.12. CONCLUSION 139

CHAPTER 7 - CONCLUSION 140
7.1. INTRODUCTION 140
7.2. MOTIVATION 140
7.3. AIM 140
7.4. RESULTS 141
7.5. RECOMMENDATIONS 146
7.6. FURTHER RESEARCH 148
7.7. SUMMARY 148
REFERENCES 150
APPENDIX A 165
APPENDIX B 166
APPENDIX C 168
APPENDIX D 169
APPENDIX E 171
APPENDIX F 174
ABSTRACT 17576
OPSOMMING 177

LIST OF TABLES

Table 5.2.1.1: Wilcoxon post-hoc: Total time……..……. 77

Table 5.2.1.2: Mean times for the total time metric: Task 1 & 2……... 78

Table 5.2.2.1: Wilcoxon post-hoc: First contact time…… 78

Table 5.2.2.2: Mean times for the first contact time metric: Task 2…….……. 79

Table 5.2.3.1: Wilcoxon post-hoc: Time on target……...……. 79

Table 5.2.3.2: Mean times for the time on target metric: Task 2…….. 80

Table 5.2.4.1: Wilcoxon post-hoc: Number of overshoots…….……... 81

Table 5.2.4.2: Mean times for the number of overshoots metric: Task 1 & 2……... 81

Table 6.3.1.1: Summary of results for the post-hoc tests: Total time - Task 1 & 2…….. 87

Table 6.3.1.2: Summary of results for the post-hoc tests: Total time sessions…….. 88

Table 6.3.2.1: Summary of Wilcoxon post-hoc tests: Total time - Task 3……... 89

Table 6.3.2.2: Summary of Mauchley's sphericity test and ANOVA……... 90

Table 6.3.2.3: Summary of post-hoc tests: Total time sessions……..……... 90

Table 6.4.1.1: Wilcoxon post-hoc: Reload time…….……..……..……... 93

Table 6.4.1.2: Summary of Mauchley's sphericity test and ANOVA……..……..... 94

Table 6.4.1.3: Summary of post-hoc tests: Reload time……..……..…… 94

Table 6.5.1.1: Results of Wilcoxon tests: Initial contact times……..…… 97

Table 6.5.1.2: Results of Mauchley's sphericity test and ANOVA……..……. 98

Table 6.5.1.3: Post-hoc test results: Initial contact - Task 1…….……..……... 98

Table 6.5.2.1: Wilcoxon post-hoc: First contact time…….……..…….…… 100

Table 6.5.2.2: Results of Mauchley's sphericity test and ANOVA…….…….. 101

Table 6.5.2.3: Post hoc results: First contact time - Task 2…….……...…… 101

Table 6.5.3.1: Post-hoc tests: Initial contact times…….……..……. 103

Table 6.5.3.2: Results of Mauchley's sphericity test and ANOVA……... 103

Table 6.5.3.3: Post-hoc test results: Task 3……...…….……..…….. 104

Table 6.6.1.1: Summary of Wilcoxon post-hoc tests: Number of overshoots…..……. 107

Table 6.6.1.2: Summary of Mauchley's sphericity test and ANOVA…..…….…… 108

Table 6.6.1.3: Summary of post-hoc tests: Number of overshoots…..…….……. 108

Table 6.6.2.1: Summary of Wilcoxon post-hoc tests: Number of overshoots…..……. 110

Table 6.6.2.2: Summary of Mauchley's sphericity test and ANOVA…..…….…… 111

Table 6.6.2.3: Summary of post-hoc tests: Number of overshoots…..…….…… 111

Table 6.6.3.1: Summary of Wilcoxon post-hoc tests: Number of overshoots…..……. 114

Table 6.6.3.2: Summary of Mauchley's sphericity test and ANOVA…..…….……... 114

Table 6.6.3.3: Summary of post-hoc tests: Number of overshoots…..…….……... 115

Table 6.7.1.1: Summary of post-hoc tests: Time on since overshoot…..…….…….... 117

Table 6.7.1.2: Post-hoc tests: Time on target since overshoot…..…….…….....……. 119

Table 6.7.2.1: Wilcoxon post-hoc: Time on target…..…….…….....…….……... 120

Table 6.7.2.2: Results of Mauchley's sphericity test and ANOVA…..…….…….……... 120

Table 6.7.2.3: Post-hoc test results: Time on target after overshoot for BCIF…..……..... 120

Table 6.7.3.1: Wilcoxon post-hoc: Time on target…..……....……....……....…… 122

Table 6.7.3.2: Results of Mauchley's sphericity test and ANOVA…..…....……....…… 122

Table 6.7.3.3: Post-hoc test results: Time on target after overshoot for BCIF…..…....…… 123

Table 6.8.1.1: Summary of post-hoc tests: Number of misses…..…....……....…....…... 126

Table 6.8.1.2: Summary of post-hoc tests: Number of misses sessions…....…….....…... 127

Table 6.8.2.1: Summary of Wilcoxon post-hoc tests: Number of misses……..…....…… 129

Table 6.8.2.2: Summary of Mauchley's sphericity test and ANOVA……..…....……. 129

Table 6.8.2.3: Summary of post-hoc tests: Number of misses……..…....……... 130

Table 6.9.1: Number of non-eliminated targets……..…....……...…....…….. 132

Table 6.10.1: Post-hoc tests: Number of resets…..……..……..…....…….. 134

Table 6.11.1.1: Means and standard deviation of session 1 and 5 for BCIG and BCIF…..... 136

Table 6.11.1.2: Results of Wilcoxon tests…....……..…....……..…....……..... 136

Table 6.11.1.3: Means and standard deviation for device assessment questionnaire…..... 139

Table 6.11.1.4: Results of Wilcoxon tests…..……..…....……..……..…....…….. 139

LIST OF FIGURES

Figure 2.2.1: Emotiv BCI headset…….. 23

Figure 2.2.2: Neurosky Mindwave headset……... 24

Figure 2.2.3: The BCI electrode positions on the head…….. 24

Figure 2.3.1: CyberGlove……... 44

Figure 2.3.2: DG5 Vhand glove 3.0……... 45

Figure 2.3.3: The Peregrine gaming glove…… 45

Figure 3.1: Nielsen's Usability Model…….…….. 55

Figure 4.1: Duck Hunter…….……... 63

Figure 4.2: Emotive Brain Computer Interface……..…… 65

Figure 4.3: Emotiv mouse emulator……..……..……. 66

Figure 4.4: EmoKey application…… 66

Figure 4.5: The Peregrine gaming glove…… 67

Figure 4.6: Touch and activator points on Peregrine glove……... 67

LIST OF CHARTS

Chart 6.3.1.1: Mean times for task 1 ……..…….…….. 88

Chart 6.3.1.2: Mean times for task 2……...……...……... 88

Chart 6.3.2.1: Mean times to complete task 3……...……..... 91

Chart 6.4.1.1: Mean times for reload.……...……...…….. 95

Chart 6.5.1.1: Mean times for task 1.……..……...……... 99

Chart 6.5.2.1: Mean times for task 2.……...……...…….. 102

Chart 6.5.3.1: Mean times for task 3.……...……...……... 104

Chart 6.6.1.1: Mean number of overshoots.……...……...…..... 109

Chart 6.6.2.1: Mean number of overshoot……...……....……..... 112

Chart 6.6.3.1: Mean number of overshoots for task 3……...……....……..... 115

Chart 6.7.1.1: Mean times for task 1……...……....…….....……..... 118

Chart 6.7.2.1: Mean times for task 2……...……....…….....……..... 121

Chart 6.7.3.1: Mean times for task 3….....…...……...…...……...…...……...…...……..... 123

Chart 6.8.1.1: Mean number of misses for task 1……...……...…...……..... 127

Chart 6.8.1.2: Mean number of misses for task 2……...……...…...……..... 128

Chart 6.8.2.1: Mean number of misses for task 3……...……...…...……..... 130

Chart 6.10.2.1: Mean number of resets for task 1……...……...…...……..... 134

Chart 6.10.2.2: Mean number of resets for task 2……...…...…...……..... 135

Chart 6.10.2.3: Mean number of resets for task 3……...…...…...……..... 135

CHAPTER 1 - Introduction

1.1. Introduction

In the field of human-machine interaction, the user interface is where interaction between the user and the machine takes place. The human–machine interface can be described as the channel of communication between the user and the computer. The goal of this interaction is to allow the user to easily and effectively operate the computer (Johannsen, 2009) – or more specifically, the interface should be usable (Camara, 2011).

The traditional interaction techniques consist of a mouse and keyboard through which the user instructs the computer (Reimer, 2005). Recent developments in the way users interact with computers and input information have resulted in a gradual shift towards interfaces incorporating voice, gestures and movement. These interfaces, called natural user interfaces (NUIs) (Camara, 2011), simplify and create a more natural interaction between the user and the computer.

For the purposes of this study, the feasibility, in terms of usability, of NUIs for games was investigated. This chapter will discuss the aim, problem statement, motivation, hypotheses as well as the methodology that was implemented to test the specific hypotheses. It will also include a discussion on the scope, limitations and contributions of this study.

1.2. Aim

Literature indicates that there are many types of interaction between users and devices. For example, gestures (cf. Alejandro & Nicu, 2007; Manresa, Perales, Mas & Varona, 2005) and speech (cf. Sporka, Kurniawan, Mahmud & Slavik, 2006) have both proven to be useful ways of interaction. The aim of this study was to determine the usability of a NUI compared to a traditional bimanual interface, namely the use of a keyboard and mouse, when used within a gaming environment.

1.3. Problem statement

Although a large proportion of game console users have started using NUIs for gaming purposes (Oikonomidis, Kyriazis & Argyros, 2011), little to no change has been seen in the personal computer (PC) gaming scene (da Silva, Nogueira & Rodrigues, 2014). PC users are still using the traditional keyboard and mouse, which have been known to cause repetitive strain

injury (RSI). RSI is caused by the unnatural posture imposed on the user when making use of a keyboard and mouse. RSI may cause wrist ulnar deviation, forearm pronation, wrist extension, and upper arm and shoulder abduction (Swanson, Galinsky, Cole, Pan & Sauter, 1997). It has been found that children experience discomfort as a result of the average time spent on computers and electronic games and their body positions while using a computer, (Ramos, James & Bear-Lehman, 2005). This discomfort, associated with RSI, includes wrist, upper arm, back and neck pain (Burke & Pepper, 2002). NUIs could potentially alleviate these problems and provide more enjoyment for PC gamers. The question then arises: why has the trend seen in console-based gaming not migrated to the PC gaming scene? The possibility exists that the answer could be found by testing the usability of available NUIs when applied to PC gaming and comparing them to the traditional keyboard and mouse combination. This could reveal the cause of the absence of widespread NUI use in the PC gaming scene.

1.4. Motivation

More than 20 years ago an interesting question was posed by Nielsen (1993), namely: when should regular use of next-generation user interfaces outside of research laboratories be expected? To a small extent, selected next-generation characteristics can already be found in a number of present day PC operating systems and applications. In the gaming world, particularly in the console domain, technology is available for consoles such as the Sony PlayStation® and Nintendo Wii® where users must hold and make use of a motion controller to issue commands. User interfaces where users do not need to touch a controller in order to communicate with the machine are also already in use. For example, by using the Kinect®, a motion sensing input device manufactured by Microsoft for the Xbox 360® game console, users are able to communicate using gestures and spoken commands (Oikonomidis, Kyriazis & Argyros, 2011).

However, users playing games on their desktop computers still use the keyboard and mouse input combination (Beckhaus, Blom & Haringer, 2005). The first keyboard and mouse prototype was demonstrated in 1968 and formed part of the On-Line System, which was designed by Douglas Engelbart (Reimer, 2005). Although there has been a change from command line interfaces (CLIs) where access could only be achieved by using a keyboard, to graphical user interfaces (GUIs) where windows, icons and menus are used which can be manipulated by a keyboard and mouse, little change has been seen in the way people interact with their computers (Baecker, Grudin, Buxton & Greenberg, 2014). Many people believe that there has been little change in the basic WIMP (windows, mouse, icons and pointer) concept

(Cechanowicz & Gutwin, 2009). Therefore, it is probably more accurate to think of the GUI as a slow evolution towards an ideal interface. NUIs are perhaps the next step in this evolution as traditional input devices appear to have stagnated (Reimer, 2005).

As mentioned earlier, RSI has been identified as one of the drawbacks of traditional user interfaces. This raises the question as to whether there are better options available in terms of user interfaces that can successfully replace the keyboard and mouse combination. Thus, the need arises for more natural user interfaces where the user can directly interact with the computer without having to perform repetitive actions that are not natural, but have to be learned first.

The NUIs that were investigated used the Emotiv brain computer interface (BCI) as well as the Peregrine gaming glove. BCI is a fast-growing field and has added a new dimension of functionality to human-computer interaction (HCI). BCI devices have created a new channel of communication, especially for those who are unable to generate the necessary muscular movements in order to use a typical HCI device (Vallabhaneni, Wang & He, 2005). The product gives you the opportunity to influence a virtual world with your feelings, thoughts and expressions. Emotiv is a neuro-engineering company that claims users can use the Emotiv EPOC to connect to current PC games and interact with them in a completely new way. They have also developed games specifically designed for the Emotiv EPOC (Emotiv Inc, 2015).

The Peregrine gaming glove, developed by Brent Baier, is a glove that replaces a keyboard for PC games and can also be used as an input device for normal computer use. It has 17 touch points and 3 activator pads that allow for over 30 user-programmable actions. By simply touching your fingertips with your thumb you can control your game or issue commands while working on the computer. Software distributed with the glove allows you to create almost any custom combination of keys (Peregrine, 2013). The Peregrine gaming glove can be used in conjunction with a mouse or, as in this study, with the Emotiv's accelerometer. The head-mounted accelerometer was used to move the cursor, while the glove was used to issue commands.

1.5. Research question

In order to address the aim of this study, the following question was answered to determine the usability of a NUI compared to a traditional bimanual interface within a gaming environment:

To what extent is the usability of a game influenced by using a NUI as opposed to a traditional keyboard and mouse combination?

1.5. Hypotheses

A simple research hypothesis evaluates a relationship between two or more groups (Creswell, 2013). The groups in this study comprise three different types of user interfaces. The general hypothesis for this study states: There is no difference between the usability of a NUI and GUI when used for gaming purposes. Since usability consists of a number of components, secondary hypotheses were formulated to address each of these components. These are:

- $H_{0,1}$: The interface used has no effect on the effectiveness of task completion in a 2D game.
- $H_{0,2}$: The interface used has no effect on the efficiency of task completion in a 2 two-dimensional (2D) game.
- $H_{0,3}$: There is no difference between the learnability of a NUI and a GUI.
- $H_{0,4}$: The interface used has no effect on the level of satisfaction of the user in a game.

Specific hypotheses, as they relate to the metrics being evaluated, will be stated during analysis of the data.

1.6. Methodology

As discussed in Chapter 3.5.1, for this study, usability refers to the extent to which a product can be used by a specific user to achieve a specified goal with effectiveness, efficiency and satisfaction, while giving the user specific training and support (ISO 9241, 2000). Nielsen (2012) defines usability as having 5 quality components, namely effectiveness, efficiency, learnability, memorability and satisfaction. For the scope of this study, the components that were tested are effectiveness, efficiency, learnability and satisfaction. Due to time constraints memorability was not investigated. In order to test three of the four stated hypotheses, usability testing, specifically controlled user testing in a laboratory, was conducted on all proposed NUIs. Questionnaires were used to test the last hypothesis that relates to user satisfaction.

The traditional interface (in conjunction with the NUIs) was included to facilitate comparison between the interfaces. The game that was used during the user testing is a 2D shooter type game that had been developed for this study using XNA framework 4.0, and is based on a game

called Duck Hunt that was developed by Nintendo (Nintendo, 2014). Data was gathered by requesting participants to perform certain gaming tasks while making use of each interface. User tests included tasks that measured effectiveness, efficiency and learnability. Effectiveness and efficiency were measured by capturing metrics specific to these usability components. For example, the percentage of tasks completed can be used to determine the effectiveness, while the time taken to successfully complete a task is a measure of efficiency. Participants were required to complete multiple sessions in order to accurately measure learnability. Differences in afore-mentioned metrics over time were used as an indicator of learning. This quantitative data was supplemented with questionnaires to measure subjective satisfaction. Tasks were specific to the game being tested and included objectives such as shooting a required number of targets within the time allotted. Actual metrics will be discussed in Section 4.4.

Data was analysed using appropriate statistics (IBM's statistical analysis software called SPSS was used) to determine whether the stated hypothesis was valid. Descriptive statistics was applied to describe the data. The data was tested for normality, where after inferential statistics, as determined by the particular data set, was utilised. The tests used will be discussed where applicable. The analysis of the metrics will be discussed in Chapter 6.

1.7. Scope of the study

The focus of this study was to analyse the difference in usability between NUIs and GUIs during gameplay (if any). During this study one GUI and two NUIs was investigated. The GUI was represented by the keyboard and mouse, while the NUIs comprised of two combinations of the Emotiv BCI and the Peregrine gaming glove.

No three-dimensional (3D) games were used for this study. A 2D-shooter game called Duck Hunter was used as a tool to gather usability data on the various interfaces.

1.8. Contribution

The use of NUIs in the PC gaming field is low, contrary to that in the area of console gaming. Additionally, RSI has been attributed to the use of the keyboard and mouse combination. It is therefore strange that NUIs have not been accepted more widely. This investigation contributes to the body of knowledge by attempting to discover the reasons behind this phenomenon. To date, there has been little research in this particular area. This study addresses these

shortcomings as well as contributes to the body of knowledge within the field of NUIs, specifically in the area of gaming.

The knowledge gained from this study may explain why users feel that there has been little change in the way they interact with their computers (Reimer, 2005). The following topics were investigated during this study in an attempt to reveal the reasons behind the slow adoption of NUIs in the area of computer gaming:

- The NUIs may not be as effective and efficient as the traditional keyboard and mouse combination.
- There may be an element of learning that is required before these NUIs can be successfully used, suggesting that they are not as intuitive as initially thought.
- The users' level of satisfaction with the NUIs does not allow for the replacement of the traditional interface combination.

Investigation of these aspects may reveal why the adoption of the NUI has been slow. The study posed to reveal benefits or difficulties that users may experience while making use of these NUIs, which would have to be addressed before widespread adoption may take place.

This investigation also revealed what the influence of the combination of modalities were on the effectiveness and efficiency of the interface combination.

1.9. Limitations

The NUIs were selected based on availability to the researcher. There are other and newer NUIs available but due to financial constraints they were not used during this study. The NUIs that are currently unavailable to the researcher can be investigated in future studies as they become more readily available.

Three tasks that were specifically developed for this study were used during user testing, and not during actual gameplay. Actual gameplay could introduce elements not present in the tasks which could affect usability. No 3D-games were used for this study. If the current study yields positive results, future studies can be undertaken with the focus on NUI use in a 3D gaming environment.

Only four of the five quality components as described by Nielsen (2012) formed part of this study to test the usability of the user interfaces. Due to time constraints the memorability

component was not tested as it required the study to be prolonged. In order to test the memorability component, enough time has to pass to determine if users can remember how to operate the user interface after not being in contact with it for a period of time.

Although the natural user interfaces investigated during this study are very flexible and can be customised to suit individual preferences, a certain configuration was set up prior to testing and was then used to test every participant using the same test parameters.

For the scope of this study eighteen participants were selected to allow the researcher to gain enough insight into the technology being investigated to compare the different user interfaces and render an accurate verdict. The small sample size was due to the fact that participants had to be tested repeatedly on an individual basis, making use of each interface. The disadvantage of using a small sample size is that generalisation to a broader population becomes difficult. However, as stated by Nielsen & Landauer (1993), in usability testing five participants are sufficient to gain insights. As you add more participants, you will gain fewer new insights because you will keep seeing the same results being repeated. Thus, by starting with 18 participants it should allow the researcher to gather adequate data to test the hypothesis. Furthermore, only five user testing sessions were conducted for the same reason, where more testing sessions may reveal additional information that was not revealed in the first five.

1.10. Outline of the dissertation

The dissertation is outlined as follows:

Chapter 1: Chapter 1 provides an introduction to the study. Furthermore, the aims, motivation, hypotheses, methodology, scope, contribution and limitations identified are discussed.

Chapter 2: Chapter 2 will include a broad discussion on the literature which formulated the study. This will provide motivation for the study and the interaction techniques used. Therefore, the chapter will discuss human-computer interaction, brain computer interface technologies as well as glove-based technologies. Additionally, previous studies which relate to the current study will be discussed.

Chapter 3: In Chapter 3 the research and experimental design used in this study will be discussed. This will include a discussion on the methodology applicable to the current study, focusing on the experimental research design, usability testing, sampling as well as surveys.

Chapter 4: The methodology used in this study will be discussed in Chapter 4. This discussion will include the specifics concerning the game used, the specific tasks that will form part of user testing as well as the metrics that will be evaluated.

Chapter 5: The results of the pilot study will be discussed in Chapter 5. The pilot study was conducted to determine the optimal sensitivity setting to use for the head-mounted mouse, as well as to verify the research process.

Chapter 6: Chapter 6 will include the analysis and results of the study. The data for all three interfaces and tasks will be analysed and discussed. Descriptive and inferential statistical tests were applied to the data and the results thereof will be the focus of this chapter. Results as they relate to previous studies and the implications thereof will be discussed also.

Chapter 7: Chapter 7 will conclude the study by summarising the results, highlighting future research prospects and discussing the contribution of the study to the field of HCI.

1.11. Summary

The purpose of this study was to determine the reasons behind the limited use of NUIs in computer gameplay. The usability of two NUI combinations was compared to that of the traditional keyboard and mouse combination. This may lead to results and observations that can clarify the absence of widespread NUIs adoption in the PC gaming environment. This chapter discussed the problem statement, aim and motivation of the study. It also included the methodology that was applied, the scope of the study as well as the contribution and limitations of the study.

The next chapter will focus on the literature that motivated this study.

Chapter 2 – Literature Review

Introduction

This chapter will include a broad discussion on the literature which formulated the study. This will provide motivation for the study and the interaction techniques used. Therefore, this chapter will discuss human-computer interaction, brain computer interface technologies as well as glove-based technologies. Additionally, previous studies which relate to the current study will be discussed.

2.1. Human-Computer Interaction

Human-computer interaction (HCI), also referred to as man-machine interaction or interfacing, originated from the intertwined areas of computer graphics, operating systems, human factors, ergonomics, industrial engineering, cognitive psychology and computer science systems (Hewett, et al., 1996; Lazar, Feng & Hochheiser, 2010). The need for HCI as a field arose from the fact that sophisticated machines are of no value unless they can be properly operated by individuals (Karray, Alemzadeh, Saleh & Arab, 2008).

HCI can be defined as a discipline that focuses on the design, evaluation and implementation of interactive computing systems for use by humans, including the investigations of the phenomena surrounding these actions (Hewett, et al., 1996). HCI can also be defined as the study of the topics that result from people encountering computer-based technologies, and how the understanding of these topics can be utilised to improve the design of new technologies (Hooper & Dix, 2012).

From a computer science perspective, the focus of HCI is on interaction between one or more users and one or more machines. Since the meaning of the terms interaction, user and machine are sometimes ambiguous, this leads to a wide range of possible research topics in the field of HCI (Hewett, et al., 1996). For the purposes of this study HCI will be defined as a field that focuses on the design, evaluation and implementation of interactive computing systems for use by humans, with particular focus on the evaluation of interactive computing systems. One area of HCI, namely usability testing, focuses on the evaluation of the interaction between man and machine.

This study can be classified as a HCI research experiment as the usability of several input devices used for gameplay was evaluated. As stated in the definition of HCI, the field of HCI includes the evaluation of interactive computing systems for human use. Thus, the evaluation of user interfaces for human use during gameplay can be classified as an HCI study. In order to evaluate interfaces they first need to be understood, thus the interaction between the user and the computer, known as interfacing, will be discussed in the following section.

2.1.1. Interfacing

The principal task of computer input is to transfer information from the user's brain to the computer (Jacob, 1996). Thus, input devices have to conform to the user's anatomic, biomechanical, perceptual, and cognitive needs and capabilities (Taveira & Choi, 2009). Therefore, in order to operate and instruct a computer, input devices need to detect and communicate the user's physical properties to the computer. These properties may include various bodily actions performed by the user to convey his/her intent. These properties are then converted into predefined signals that are used to communicate the user's intentions to the computer (Taveira & Choi, 2009).

Interfacing has been an area of interest for as long as computers have been part of everyday life. The interaction methods that humans used to operate computers have come a long way since the introduction of computers. Computer input once consisted of actions such as actuating switches and knobs and plugging and unplugging wires. For many years after that, the primary form of computer input was by means of a punch card. Users punched the input information as holes in paper cards that could be read by the computer. After punch cards, a device called the Teletype, with a keyboard similar to that of a typewriter, was used. This allowed characters to be typed and transmitted directly to the computer. Terminals, keyboards and displays formed the basis of computer input for many years (Jacob, 1996) and are still popular today. New technologies and systems are continuously emerging and the research interest in this area has been growing rapidly in the last few decades (Karray, et al., 2008).

Interfaces are divided into categories based on the human ability they make use of to transfer the user's intent, as well as the variety of communication channels they combine, for example vision and speech. The following section will discuss a selection of the diverse interface technologies available, some much more successful and widespread than others. These

different interface technologies can be categorised into different groups by looking at the mode of input that they utilise.

2.1.2. Modes of input

The configuration of an HCI device is the most significant feature of its design. An interface primarily relies on the quantity and variety of its communication channels, which facilitate interaction between the user and computer. Each of the different channels is known as a modality (Karray, et al., 2008). An interface is defined by the number and range of inputs and outputs it offers and could be classified as either unimodal or multimodal. For the sake of brevity only the modes of input essential to this study will be discussed in detail.

2.2. Unimodal Interface Systems

A system that operates on one modality is called unimodal (Karray, et al., 2008). Categorised by the nature of the different modalities, the interfaces can be separated into three groups, namely visual-based, audio-based and sensor-based HCI. The following sections will shortly describe each group.

- **Visual-Based HCI**

The visual-based HCI area is the most established area in HCI research and is also the most commonly used input technology. Researchers focused on different aspects of human responses which can be accepted as a visual signal and can include either switch-based or pointing devices. Switch-based devices include any kind of interface that uses buttons and switches, for example, a keyboard. Examples of pointing devices are mice, joysticks, touch screen panels, graphic tablets, trackballs, and pen-based input (Karray, et al., 2008).

Some of the main research areas are facial expressions (by using video analysis) (cf. Savva & Bianchi-Berthouze, 2012), body (cf. Helten, et al., 2013), hand gestures (cf. Donovan & Brereton, 2005) and gaze detection (cf. MacKenzie, 2012).

- **Audio-Based HCI**

Another important unimodal type of HCI systems involves audio-based interaction between a human and a computer. These are devices that use speech as input, and usually need some kind of speech recognition software to evaluate spoken instructions. In order to perform this task the software makes use of its ability to be programmed to recognise a user's commands. The majority of mainstream computers have basic speech recognition built into their operating

systems (Jacob, 1996). Recent studies have shown that speech recognition can be used successfully as an alternative means of input (cf. Vasantrao, Prakash, Prakash & Anant, 2014; Miao, Metze & Rawat, 2013). Speech recognition was not included in this study, although it may later be combined with the technology utilised during this investigation.

- **Sensor-Based HCI**

This interface category is a combination of a range of HCI areas. The shared characteristic of these different areas is that at least one physical sensor is used to offer communication opportunities between users and computers (Karray, et al., 2008).

a) Use of the hands

The standard keyboard design has been difficult to displace as the primary means of computer input due to its global popularity and rather inexpensive production costs. In recent years the main driving force behind changing the standard input method was the change in the size of computers – smaller computers call for smaller, more intuitive input devices. The typewriter keyboard has become the largest component of most computers and hand-held devices (Jacob, 1996). Therefore, it has become the one element standing in the way of reducing the overall size of the computers, as the majority of laptop and desktop keyboards follow the international standards of 19 mm key spacing (Pereira, Laroche, Rempel & Hsieh, 2014).

Used in conjunction with workstations or portable computers, keyboards are one of the primary input devices currently in use. The standard configuration of the keyboard is derived from that of typewriters and, with the addition of function keys and a numerical keypad, remained mostly unchanged from the 1968 design (Lewis, Potosnak & Magyar, 1997). An alternative design, introduced to reduce stress on the hand and wrist, preserves the key layout of previous keyboards but changes the geometric arrangement of the keys. These altered keyboards are normally divided into two sections, one section for each hand. These sections are rotated away from each other to better fit the natural position of the user's hands (Jacob, 1996) and thereby alleviate stress on the wrist joint. The interfaces suggested in this study may assist in the reduction of RSI, due to glove and BCI technology being included that may be more natural to use and thus result in less stress to the user's wrist.

The fact that the traditional keyboard has been difficult to replace as a widespread interaction device has been an important motivation for conducting this study. Therefore, several natural user interface combinations was evaluated and compared to the traditional combination. These

natural user interface combinations included BCI and glove-based input devices. This investigation was conducted in an effort to conclude whether the usability of the NUIs is the reason why there has not been a shift towards natural interfaces in everyday computer use.

b) Gloves

Glove input technology detects the configuration of the fingers of the user's hand. This is referred to as a hand posture in contrast to a hand gesture. Glove technology can make use of optical fibres, which diminishes light when bent, to detect the configuration of the hand. Other glove technologies make use of mechanical sensors to detect the hand's configuration. Both of these types of devices detect the bend angle of each of the joints of each finger. Some gloves also detect the angles formed by the separation of the fingers from each other (Jacob, 1996). A third category of glove uses touch sensitive areas on the glove to detect finger presses and uses this information as input (Shin & Hong, 2005). Glove-based technology was evaluated during this study and compared to the traditional keyboard with regard to command activation. Glove interaction technologies will be reviewed in Section 2.3.

c) Use of the head

A head tracking device is a type of alternative input device that allows an individual to control the computer by moving his/her head. A 3D accelerometer attached to the head can be used to detect head movement, which in turn can be used to control cursor position (Bérard, 1999). Another use of head movement is to perform a panning and zooming function which is similar to the use of head movement in the natural world (Hix, Templeman & Jacob, 1995). Head movement as a method of input will be discussed in detail in Section 2.2.7.6

d) Electromyography

Humans use a complex set of skeletal muscles and connecting tendons and bones to generate movement of all parts of the body. Movement of the body originates in the brain by communicating an electrical signal via the nervous system. The fibres that make up all muscles are stimulated by this signal and these fibres then contract in response to create the desired movement (Saponas, et al., 2009).

Electromyography (EMG) detects this muscular movement by evaluating the electrical potential between pairs of electrodes. This can be done invasively from the surface of the skin or by placing needles in the muscle (Saponas, et al., 2009). These signals can also be detected and measured before the muscle activity starts. This helps to reduce the overall latency of the

system (Jacob, 1996). By making use of a BCI device, facial expressions can be detected and used as input (Heger, Putze & Schultz, 2011). Investigating a BCI as a main input method for gaming formed part of this study, especially the use of facial expressions, thus EMG will be further discussed in Section 2.2.2.3

e) Electroencephalogram

Electroencephalogram (EEG) signals are important sources of data to evaluate the brain processes that form thoughts and actions. EEG signals that are produced during a mental task can be classified and then detected as input (Lebedev & Nicolelis, 2006). Therefore, EEG signals have recently been used to convey a user's intention to a computer. The use of BCI technology as an alternative method of input for gaming forms an integral part of this study. The use of thought as input seems to be a potential input option for gaming, thus EEG will be discussed in Section 2.2.2.1.

Multimodal Interface Systems

Modalities refer to the different senses that individuals utilise, for example sight, hearing, touch, smell and speech. Multimodality is a combination of different modalities. Human interaction with the world is naturally multimodal as individuals make use of numerous senses in order to passively and actively interact with their immediate environment (Turk, 2014). Due to the nature of human communication, multimodal systems are fundamentally different to standard GUIs. As a result, basic tasks in a GUI become more complicated to interpret as these actions now require recognition and are thus susceptible to misinterpretation. Multimodal interfaces require the computer to process two or more input streams that are delivered simultaneously (Oviatt & Cohen, 2000).

The number of input modes, their types and how they operate together may differ between multimodal systems. Diverse combinations of gesture, speech, facial expressions, gaze and other means of input are included in multimodal interfaces. The combination of gesture and speech (cf. Hoste, Dumas & Signer, 2012; Kristensson & Vertanen, 2011) is a commonly supported multimodal interface. A positive effect of multimodality is that the cooperation of different modalities can assist with recognition. For example, visual-based lip movement detection can help with speech recognition, which is audio-based, whereas speech recognition can assist in gesture recognition, which is visual-based (Karray, et al., 2008).

During the current study different multimodal interfaces were tested as BCI and glove technology were combined to form two separate multimodal input combinations. The motivation behind this was to determine if possible disadvantages of one modality can be negated by using another modality. If no such disadvantages were present, this approach would determine whether a multimodal interface is more usable than a unimodal interface in such an instance. These combinations were evaluated in order to compare them to the traditional keyboard and mouse interface in terms of usability. Traditional and natural interface technologies will be discussed in the following sections.

2.1.3. Traditional Interfaces

Traditionally, a user operates a computer with a standard point and click mouse and a so-called QWERTY (the top left-hand configuration of keys) keyboard (Taveira & Choi, 2009). The QWERTY keyboard is the oldest and most common computer input device. Some of the traditional QWERTY keyboard's limitations were identified during the 1920s. These limitations relate to the long travel distances between keys, the necessity to use weaker and less agile fingers to make full use of the keyboard, as well as making use of the less dominant hand for interaction (Swanson, Galinsky, Cole, Pan & Sauter, 1997). Additionally, the user's dominant hand must switch between the keyboard and mouse for input purposes (pointing and clicking actions), which is often a considerable distance (Jacob, 1996) thus resulting in a travel time limitation to this combination of input devices.

Injuries such as wrist ulnar deviation, forearm pronation, wrist extension, and upper arm and shoulder abduction are associated with the use of the traditional keyboard (cf. Marklin, et al., 1999; Swanson, et al., 1997). The exaggerated or unnatural posture characteristics of making use of a keyboard have received continuous research attention which resulted in the development of multiple alternative input technologies and devices (Taveira & Choi, 2009).

With current user interface technology the amount of information that is communicated from the computer to the user is far greater than the amount of information traveling from the user to the computer. Graphics, audio and other media can output large amounts of information, but there are no methods for inputting the same amounts of information from the user. Because of the users' abilities, they can receive vast amounts of information but are not capable of generating information to such a degree. Communication between computers and users is therefore mostly one-sided (Jacob, 1996).

Researchers are interested in user interface technology that can help redress this imbalance by procuring input data from the user in a convenient and swift manner (Jacob, Leggett, Myers & Pausch, 1993). Additional modes of communication could be the key to this objective (Tufte, 1989). Traditional computer input depends on physical interaction with control devices. In an attempt to make it less cumbersome for the user to communicate his intentions, companies and research groups are very interested in more natural methods of interaction (Plass-Oude Bos, et al., 2010).

Therefore, this study focussed on comparing different combinations of natural user interface technologies to the traditional user interface in order to find a solution to the above-mentioned problems, which includes the injuries associated with keyboard use as well as the lack of input produced by users. This entails the investigation of more natural means of input that may allow the user to communicate larger amounts of information to the computer by making use of an alternative combination of modalities.

2.1.4. Natural user interfaces

As illustrated in the previous section, existing physical technologies for HCI can be categorised by the human senses and abilities that the device utilises for input. These devices rely on human senses that include, but are not limited to, vision, speech, thought and touch (Karray, et al., 2008). Recent technologies in HCI are now combining prior methods of interaction with new advances in interaction technology. Some of these new devices upgraded and integrated preceding devices to form new interaction technologies. With technology improving at the current rate, the borders between these new technologies are also closing fast (Karray, et al., 2008).

Due to the afore-mentioned problems associated with the traditional input combination of the keyboard and mouse, research is being conducted to design and test more natural and intuitive interfaces, known as Natural User Interfaces (NUIs). A NUI is a product for HCI that the user operates through intuitive and natural actions related to everyday behaviour. This means that the user is able to use the interface with little or no training and that the user should enjoy using the interface (Steinberg, 2012). If users are able to interact with the virtual world in a way that is similar to the real world, then learnability and memorability will no longer pose a problem (Plass-Oude Bos, et al., 2010). In order to achieve this, the restrictions have to be removed from the communication channel, resulting in faster, more natural and more convenient means

for users and computers to exchange information (Turk, 2014). On the user's side, the constraints are in the nature of the human's abilities while on the side of the computer, the only constraint is the range of devices that can be designed (Jacob, 1993).

This new notion of input requires new technologies, interaction techniques and software to deal with them. NUI researchers focus on developing new technologies, such as interaction methods (including interfaces) in order to remove current restrictions on what is possible in HCI. The main objective of this research is to make full use of human communication and interaction capabilities when interacting with computers (Turk, 2014).

Thus, the goal of natural interfaces is to increase the flow of information between the user and the computer. This will result in user-computer interaction that is more comparable to the user's everyday tasks and behaviour (Jacob, 1996). There are different types of natural user interfaces, such as touch screens, gesture recognition, speech recognition and BCIs, as well as the tracking of eye movement.

In seeking natural modes of interaction, the user's input actions should be as close as possible to the user's thoughts that stimulated those actions. Thus, the user interface must embody the user's abilities, meet the user's needs as well as take full advantage of his/her capabilities (Wigdor & Wixon, 2011). This could potentially bridge the gap between the user's intentions and the actions necessary to carry out these intentions (Lee, Isenberg, Riche & Carpendale, 2012). The impetus for this strategy is that it builds on the behaviour and skills that humans have attained through evolution and experience. When utilising natural interfaces users should be able to interact with technology using the same gestures used to interact with objects in everyday life. Rather than training a user to operate a keyboard or mouse, the full range of human senses can be used to leverage existing, natural communicative abilities for input and output devices (Jain, Lund & Wixon, 2011).

Since human interaction with the real world is multimodal, this type of interaction is part of what defines a natural experience on the part of the user (Jain, Lund & Wixon, 2011). Therefore, during this study multimodal input was investigated by making use of a combination of glove-based and BCI input. However, many users might differ in the way they interact with technology, consequently not all users experience the same method as natural (Malizia & Belucci, 2012). This may lead to some users not being able to make proper use of a specific interface as they will not experience the interaction as natural. Therefore, users will find some

natural user interfaces more enjoyable than others due to their predisposition to the specific device and mode of interaction. Consequently, the usability thereof needs to be evaluated in order to accept these technologies into mainstream use.

One important characteristic of NUIs is flexibility, which allows users to customise their interface to better suit their needs, resulting in more efficient and intuitive use (Steinberg, 2012). The natural user interfaces investigated during this study are very flexible and can be customised to suit individual preferences, for example choosing to use certain facial expressions rather than others. This, however, did not form part of the current study as a certain configuration was set up prior to testing in order to test every participant using the same test parameters. If the use of the interfaces proves to be favourable, the customisation should serve to enhance the interface and will therefore not be tested. The optimal sensitivity setting for the head-mounted mouse was investigated in the pilot study and will be discussed in Chapter 5.

The eventual aim of NUIs is to bring HCI to an endpoint where user interaction with computers will be as natural as interaction between humans (Rautaray & Agrawal, 2015). The NUIs that were evaluated during this study will be discussed in the following subdivisions.

2.2. Brain-Computer Interface (BCI)

This section will discuss BCI technology, including how the technology functions, the products available as well as the different applications of the technology in everyday life.

2.2.1. Direct Brain-Computer Interaction

For the nervous system to operate effectively, neurons have evolved distinctive abilities for communication within the cell as well as communication between cells. In order to support fast, long distance communication, neurons have developed special capabilities to send electrical signals called action potentials (Stufflebeam, 2008). These neurons are constantly firing across the brain at very fast rates (Niedermeyer & da Silva, 2005), which allows for voltage differences on the scalp to be recorded. The data received from recording these electrical signals can be used as an alternative means of input within the field of HCI as a BCI (Curran & Stokes, 2003). These electrical signals can be classified into different categories, namely electroencephalogram (EEG), electrooculography (EOG) and electromyography (EMG) signals. These signals are considered among the most important sources of

physiological information when making use of brain-computer systems (Fatourechi, Bashashati, Ward & Birch, 2007). These three types of brain signals are discussed in the following sections.

2.2.2. Brain signals

2.2.2.1. Electroencephalogram (EEG)

The first recording of this nature was made by the German psychiatrist Hans Berger in 1924 (Tudor, Tudor & Tudor, 2005). These electrical signals are referred to as electroencephalogram (EEG) signals and are a key source of information for studying the core brain processes that form human thoughts and actions (Benjamin & Keller, 2003). EEG data is particularly useful for medical diagnoses and distinguishes between an assortment of central nervous system irregularities (Quinonez, 1998), for example, identifying the source of seizures in epilepsy patients (Singer, 2008). Therefore, a routine EEG remains an inexpensive, widely used diagnostic tool in the medical field (Quinonez, 1998).

Until a few years ago, caps were used to place the sensors on the participants' head. These caps were complex to position due to the numerous wires that were connected to it. Recently this problem has been surmounted by the introduction of wireless sensor technology (Hondrou & Caridakis, 2012).

EEG recorders with up to 256 electrodes are being used, and experiments using them produce large amounts of raw data resulting from the detection of these brain rhythms. Thus, EEG can be used as a tool to image the brain while it is performing a cognitive task (Singer, 2008). If an HCI system is based on EEG data, then it is referred to as a BCI. EEG signals being contaminated by muscle activity is a well-known problem for BCI use, and these contaminants need to be detected, isolated and removed from the EEG signal to ensure proper communication. Scientists who make use of EEG attempt to filter out this noise through various means. By applying complex signal analysis they endeavour to improve sensor sensitivity to increase the quality of the EEG signal (Singer, 2008).

Originally most BCI applications primarily focused on the severely disabled (Fatourechi, Bashashati, Ward & Birch, 2007). Only recently have game developers started to focus on using EEG signals to allow players to influence and control virtual environments with their thoughts (Singer, 2008). Changes in brain rhythms such as mu, beta and gamma rhythms,

which are related to movement, are detected and can be used as input. When the user of the BCI learns how to control these rhythms, the mu or beta rhythm amplitudes can then be converted into cursor movements on a computer (Wolpaw & McFarland, 2004). The use of EEG as a possible input method for gaming was further investigated in order to see whether adequate levels of control was possible. The possibility of using this technology during the study will be discussed in Section 2.2.7.4.

2.2.2.2. Electrooculography (EOG)

A known problem with BCIs is that slight muscle movements called biological artefacts can generate electrical potentials more than 10 times as strong as those produced by neurons (Singer, 2008). These artefacts can change the appearance of neurological phenomena and can even be incorrectly used as the source data for BCI systems. One of the most well-known and important artefacts is known as electrooculography (EOG) (Fatourechi, Bashashati, Ward & Birch, 2007). The EOG signal stems from an electrical potential that is generated across the cornea and retina when individuals move their eyes (Anderer, et al., 1999), and can be measured without much difficulty (Usakli, Gurkan, Aloise, Vecchiato & Babiloni, 2010). By making use of horizontal and vertical eye movements or blinking, users can interact with computers and other electronic devices (Fatourechi, Bashashati, Ward & Birch, 2007). Therefore, EOG can be interpreted as an interfering signal in terms of EEG use or as an alternative source of interaction data on its own (Usakli, Gurkan, Aloise, Vecchiato & Babiloni, 2010). The use of blinking in order to activate a command was investigated for possible use in this study. The use of EOG will be further discussed in Section 2.2.7.5.

2.2.2.3. Electromyography (EMG)

Electrical signals generated by moving the head, jaw, tongue or body are referred to as EMG. During the generation of EEG signals, the signal is usually contaminated by EMG, which increases the difficulty of analysing the signal for the purpose of input. When a complicated task or activity is being performed it may lead to more facial movement, which will cause EMG signal spikes (Fatourechi, Bashashati, Ward & Birch, 2007). This results in increased contamination of the EEG signal when the user is focusing on the already difficult task of using the BCI device.

BCI input in the past was usually intended for use by individuals with severe motor disorders and cerebral palsy, thus removing the EMG from the brain signal is very important, as these

disorders are frequently associated with involuntary contractions of the cranial or facial muscles (Fatourechi, Bashashati, Ward & Birch, 2007). EMG can thus be a contaminant of the EEG signal but can also be utilised as a primary mode of interaction by detecting and recognising facial movements. The possibility of using facial expressions as the primary method for activating commands during this study was further investigated. The use of EMG as a primary input method will be discussed in Section 2.2.7.5.

2.2.3. BCI Technology

BCI technology makes it possible to detect these different brain signals (Singer, 2008). A BCI is a device that identifies information directly from the brain, facilitating real-time responses and acts as a communication system where commands from the user to the computer do not pass through the body's normal output, for example the nerves and muscles (Wolpaw, Birbaumer, McFarland, Pfurtscheller & Vaughan, 2002), but are rather directly communicated to the computer. A BCI is an input modality that can detect certain actions, intentions and psychological states (e.g. cognitive and emotional states, as well as facial movement) by capturing and analysing the user's brain activity (Gurkok, Nijholt & Poel, 2012). Thus, a BCI is a direct communication corridor between the brain and an external device and is therefore a potentially influential communication and control option for users (Van Erp, Lotte & Tangermann, 2012).

2.2.4. BCI Classifications

The neural activity used by a BCI can be recorded using non-invasive or invasive techniques. Non-invasive BCIs make use of EEG activity on the scalp while invasive BCIs obtain data from the cortical surface or brain (Reyes & Tosunoglu, 2011).

2.2.4.1. Non-Invasive Method

As previously mentioned, brain activity produces electrical signals that are detectable on the scalp, on the cortical surface, or within the brain. Non-invasive BCIs translate these signals from the scalp by using a set of electrodes, in the form of a headset which is placed on the user's head (Wolpaw, Birbaumer, McFarland, Pfurtscheller & Vaughan, 2002). No surgery is required for the non-invasive method. In order for a computer to process information from the brain, the EEG signals are recorded, amplified and digitised (Myeung-Sook, Joonho & Sunghoon, 2010). The drawback with this method is that the signal quality is lower than that

obtained from an invasive technique (Jimenez, Andujar & Gilbert, 2012). Furthermore, swift sequences of actions may be restricted when using non-invasive devices as EEGs reflect on slow changes in the mental state of the user (Reyes & Tosunoglu, 2011).

2.2.4.2. Invasive and partially invasive method

As previously mentioned, neurons have developed special capabilities to send electrical signals called action potentials (Stufflebeam, 2008). Invasive BCIs detect the user's intent by making use of these neuronal action potentials that are found within the cerebral cortex. The cerebral cortex is the outer layer of grey matter covering the two brain hemispheres and is typically 2 - 3 mm thick. The cerebral cortex has many functions including memory, language, abstraction, creativity, judgment, emotion and attention, as well as the creation of movement (Swenson, 2006). Therefore, invasive BCIs require direct contact with the cerebral cortex or other sections of the brain. This is accomplished by performing an invasive neurosurgery as chips or intra-cortical microelectrodes are implanted directly into the individual's brain (Lal, et al., 2005).

The advantage of this method is that it has the highest signal quality but the method is very expensive (Reyes & Tosunoglu, 2011) and also poses ethical concerns. Preclinical studies that address the risks of invasive BCI methods and establish the usefulness of the technology first need to be conducted in order to alleviate these concerns (Wolpaw, et al., 2006).

EEG signals can also be measured using a partially invasive method. This requires surgery in order to implant a chip that will rest on the scalp. The advantage is that it not only has better signal resolution than the non-invasive BCI, but there is also a smaller risk of creating scar tissue as with the invasive method. This method is also expensive and requires surgery, which is a disadvantage (Reyes & Tosunoglu, 2011).

Most BCI research is conducted using the non-invasive method due to the disadvantages associated with the invasive and partially invasive methods. The ethical predicament of invasive brain surgery on humans and animals should also be taken into account with regard to BCI studies (Wolpaw, et al., 2006).

Invasive as well as partially invasive BCI technologies were not used during this study due to the risks in terms of surgery and costs involved as well as the ethical concerns related to

invasive BCI methods. Consequently, only non-invasive, commercially available BCI technology were investigated for the purpose of this study.

2.2.5. Non-invasive commercially available BCIs

In recent years a number of companies have used medical grade EEG technology to create inexpensive BCIs. This technology has been used in toys and gaming devices. BCI products have been released by companies such as NeuroSky, Emotiv, Uncle Milton, MindGames, and Mattel (Van Erp, Lotte & Tangermann, 2012). Commercially available BCI headsets differ in the number of electrodes that are used for measurement, from one electrode for headsets utilising NeuroSky sensor technology to 14 in the Emotiv BCI headset. The more electrodes in a headset, the more sophisticated its potential functions (Singer, 2008). For the sake of brevity only the Emotiv BCI and the Neurosky headsets will be discussed.

2.2.5.1. Emotiv BCI Headset

Emotiv is a neuro-engineering company that released the Emotiv BCI (Figure 2.2.1) in 2009. The Emotiv BCI is aimed at the gaming market, and is not classified as a medical device, though a few researchers have since adopted it for a variety of applications. Instead of requiring a special gel, the electrodes of the BCI simply need to be dampened using a saline solution, which is both a disinfectant and easily accessible (Cernea, Olech, Ebert & Kerren, 2011).

Figure 2.2.1 - Emotiv BCI headset (Emotiv Inc, 2015)

However, the upcoming Emotiv Insight does not make use of any gel or saline solution but rather makes use of dry sensor technology (Emotiv Inc, 2015), which would shorten the preparation time before the device can be used as well as reduce rusting of the electrodes. This could in turn attract the acceptance and subsequent use by more individuals who would not have used the device before. The Emotiv Insight was not yet available at the time of testing and was thus not included in the study.

2.2.5.2. Neurosky

Figure 2.2.2 - Neurosky Mindwave headset

Emotiv's main competitor, NeuroSky, has taken a simpler approach to designing a commercial BCI. The headset developed by NeuroSky uses only one sensor that rests on the forehead of the user, as seen in Figure 2.2.2. The headset has a built-in processor that analyses the inbound signal before wirelessly sending the data to the computer. The headset detects two different states, strong concentration and deep relaxation (Singer, 2008). Since only two states can be used as input, the NeuroSky headset was not used for this study as more states were required for gameplay. For the purpose of this study the Emotiv BCI was utilised due to its increased functionality.

2.2.6. Emotiv BCI functionality

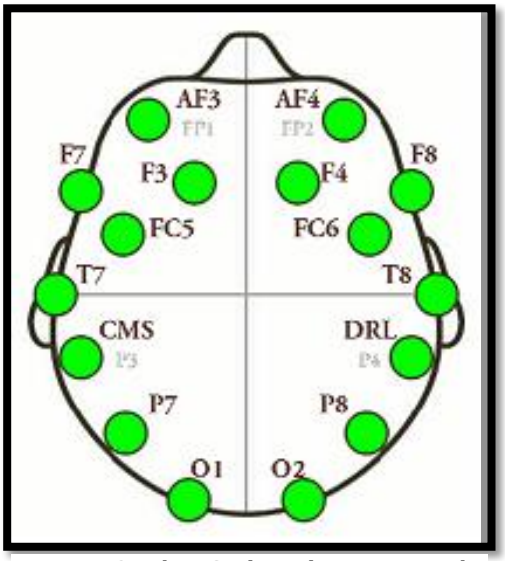

Figure 2.2.3 - The BCI electrode positions on the head.

The Emotive BCI, with its 14 electrodes (Figure 2.2.3), can monitor more complex brain states, both in terms of emotional complexity and ability to produce finely tuned control for games or machines (Singer, 2008). The Emotiv BCI has three built-in brainwave processing suites, namely Affectiv, which gives a measure of the users' subjective emotions; Cognitiv, which detects specific thoughts; and Expressiv, which detects facial expressions (Reyes & Tosunoglu, 2011). The device also has an accelerometer which is used to detect head movement (Singer, 2008).

These signals are transferred from the headset to the computer through wireless technology. A tool called EmoKey is available for mapping emotional states, thoughts, expressions and head movement to keys on the keyboard or mouse actions (Emotiv Inc, 2015). All of the above mentioned functionalities can be used through the supplied Emotiv control panel application. The panel has an interface for visualising and controlling all three suites, along with a display of the contact quality of each sensor (Emotiv Inc, 2015).

The Emotiv BCI has the ability to detect emotional states such as excitement, engagement, boredom, meditation and frustration. These can then be used as additional means of input other than the primary input device (Hondrou & Caridakis, 2012). A new generation of computer games are emerging, encouraged by the advances in sensor technology and signal processing. Several affective games encourage the user to convey his/her emotional state in order to dynamically adjust in-game actions and events to the player's emotional state. For example, the game could be made more challenging if it is detected that the player's level of boredom is increasing. In the same way, when anxiety is detected the game could slow down or decrease the difficulty level (Hondrou & Caridakis, 2012). This will increase the player's level of immersion and engagement (Kim, Bee, Wagner & Andre, 2004).

This method of interaction is a passive observation technique with little to no deliberate user action or intent. However, this is a slow method of interaction as emotional states first have to change before a new action can be triggered. Thus, making use of emotional states as a primary method of interaction will neither be effective nor efficient due to the lack of direct user control over his/her emotions. Taking the above information into consideration, the Affectiv suite of the Emotiv BCI did not form part of this study as only natural devices that could act as primary methods of interaction could be evaluated against the traditional combination of the keyboard and mouse. This technology could well be further investigated at a later stage for its alternative channel of input.

The Emotiv BCI's Cognitiv suite makes use of pre-recorded EEG signals to detect the user's intent. Thus, the Cognitiv suite enables the user to interact with the computer using only thoughts (Emotiv Inc, 2015). A major problem with cognitive input is that the brain continuously produces signals, creating difficulty in detecting whether the user is performing a task or not (van de Laar, Plass-Oude Bos, Reuderink, Poel & Nijholdt, 2013). This problem is known as the Midas touch effect, referring to the mythological figure, King Midas, who turned everything to gold by touching it. Therefore, a primary challenge in the field of BCI research is the ability to turn off neural input when the user does not wish to interact with the system (Tan & Nijholt, 2010). The same phenomenon is seen with eye tracking and methods have been developed in order to counteract it, for example, using dwell time or by using multiple modalities (Poole & Ball, 2006). The problem can therefore also be overcome when using a BCI if activation techniques are designed to counter the Midas effect. The possibility

of using cognitive input as the primary method of interaction for gaming will be further discussed in Section 2.2.7.4.

The Emotiv also has an Expressiv suite that uses muscle signals that are detected from face and eye movement allowing it to distinguish between different facial expressions. Recognising voluntary facial expressions can be achieved through analysis of the EOG and EMG data that is recorded. These facial expressions can be used as an alternative input method to control a device. While all three software suites make use of the signal data from the electrodes, the Expressiv suite is more concerned with the EMG present in the EEG signal. The use of facial expressions as a primary mode of input for games will be discussed in Section 2.2.7.5.

The headset has a built in accelerometer which can detect head movement, and this input can be used for pointing (Singer, 2008). The user can control the mouse cursor through a mouse emulator that is included in the software package. The mouse emulator uses the data sent from the accelerometer. When the user then swivels his/her head up or down, the cursor will also move up or down, and when the user swivels his/her head to the left or right, the cursor will also move to the left or right. This method of cursor control will free the user's hands, thus allowing him to make use of additional interaction devices if need be. This functionality can also aid individuals, who have lost the use of their hands, to control cursor movement. The sensitivity of the cursor movement can be changed to suit the user's needs.

During the following sections the applications of BCI technology as well as the viability of each of these discussed functions as a primary method of interaction will be investigated.

2.2.7. BCI application and functionality

2.2.7.1. Introduction

The primary focus of significant BCI research has been the investigation and development of radically new communication technologies for individuals with neuromuscular deficiencies, which prevent them from using conventional communication channels (cf. Kübler, et al., 2005; Guger, et al., 2009). Disabled BCI users are categorised into three distinctive groups, namely those who are totally locked-in, those who retain very limited levels of neuromuscular control and lastly those who still have extensive neuromuscular control (Wolpaw, et al., 2006).

Recently the research focus has shifted to improving rehabilitation rather than replacing or restoring lost abilities (Wolpaw, 2014). Several systems have successfully been used for mind-controlled games by individuals with disabilities (Heger, Putze & Schultz, 2011). Therefore, these new technologies could result in a change of the composition of BCI user groups since they are now starting to assist a wide range of disabled and able-bodied users (Allison, 2011). If able-bodied users have to use BCIs on a daily basis, BCIs must function while users move and interact with their environment, as well as require minimal training while providing robustness for accurate long term data collection (Tangermann, et al., 2009). The next section will discuss the possibility of using BCI technology for disabled and able-bodied users.

2.2.7.2. BCI use for disabled and able-bodied users

Twenty-four motor-disabled end-users were trained to use a BCI in their homes or clinics. Participants were asked to control a tele-presence robot and a text-entry system with the use of a BCI. Results showed that fifty percent of the participants were able to achieve good BCI performance as well as successfully complete the tasks (Leeb, et al., 2013). Interesting to note was the fact that disabled users performed similarly to able-bodied users who were tested in a previous study (Tonin, Leeb, Tavella, Perdikis & del Millan, 2010). Furthermore, the majority of published works in terms of BCI applications for disabled people were conducted using data from able-bodied participants (Leeb, et al., 2013). This indicates that able-bodied use of BCI technology may actually be more prevalent than with disabled individuals when taking research studies into account, thus the use of BCI technology by able-bodied users is not a new phenomenon. Several studies have shown that both disabled and able-bodied users can make use of BCI technology for communication and device control (cf. Allison & Pineda, 2003; Volosyak, Cecotti & Graser, 2009; Mugler, et al., 2010). However, due to slow speed, a high error rate and the complexity of BCI systems it has initially proven challenging for use in mainstream systems (Gurkok, Hakvoort & Poel, 2011).

Although use of BCI input by able-bodied users may not necessarily be suited to everyday tasks on a computer, for example when typing, it may be useable in the field of entertainment, for example in the gaming area.

2.2.7.3. BCI application in gaming

Since the 1990s researchers have been investigating BCI-controlled computer games. These studies were conducted with the goal of using computer games to improve participants' performance in BCI experiments as well as keeping them motivated. In recent years entertainment and gaming have become popular focus areas for BCI research (Coyle, Principe, Lotte & Nijholt, 2013).

Computer games do not take place in the real world. The same is true of our thoughts, and thus neither are constrained to what is physically possible. Therefore, expressing ourselves directly in the game world would make more sense, thereby allowing interaction to be performed without the use of physically limited body movements (Plass-Oude Bos, et al., 2010). Thus, making use of a NUI for gaming, where the interaction is more intuitive, would make sense in that the user will be less aware of the interaction method and be more immersed in the game environment. The BCI is well-suited to provide a new communication channel that requires voluntary control by the user, to the brain (Wolpaw, McFarland, Neat & Forneris, 1991). With advances in technology and research, it has become possible to process EEG data in real-time (Nicolas-Alonso & Gomez-Gil, 2012), thus increasing the applicability of these technologies in the gaming environment. As such, there has been a rising research interest in non-invasive BCI systems over the last couple of years (Heger, Putze & Schultz, 2011). The recent advances in computing and bio-sensing technologies have improved possibilities for BCI application, making them promising for use in mainstream applications (Jackson & Mappus, 2010). Since many proof-of-concept studies have shown that BCI input can be utilised to operate computers, it is reasoned that it can be utilised in the area of computer gaming (Coyle, Principe, Lotte & Nijholt, 2013).

In order to accommodate able-bodied users, preparation time and difficulty in setting up BCI devices are being reduced through the development of new sensor technology that does not require electrode gel or saline solution. This makes BCIs more accessible to mainstream users. For example, dry sensor technology can detect brain signals as well as other signals such as EOG and EMG (Allison, 2011). These new developments of affordable and wireless dry cap technology and improvement in timing, dynamics and the speed of interaction, make BCI use practical for able-bodied users (Tangermann, et al., 2009). In particular, dry electrode systems have been aggressively advertised by companies such as Emotiv and NeuroSky for gaming and other purposes (Allison, 2011). Several games have also been developed specifically for the

Emotiv BCI and can be downloaded from the Emotiv e-store. These games include Homecoming, Cortex arcade, Stonehenge, Arena as well as the Spirit Mountain demo (Emotiv Inc, 2015). Numerous innovative BCI devices and applications have recently been developed, including devices to control smart homes or other virtual environments, games, prosthetic devices, wheelchairs and robotic devices (Allison, 2011).

Currently, there are still a number of research problems that will need to be investigated in order to increase the widespread use of BCI games. These problems include eliminating the calibration process or hiding the calibration process in the gameplay, and detecting the type of BCI controls that will be the most effective for gaming purposes. Finding ways of training users in a way that does not interrupt gameplay will also form part of these investigations (Coyle, Principe, Lotte & Nijholt, 2013).

BCI technology offers a variety of input modalities, including cognitive recognition, facial expression recognition and an alternative method of cursor control (Emotiv Inc, 2015). These alternative input methods will be discussed in the following sections.

2.2.7.4. Cognitive input recognition

Thought as an input method for games has been used to move a game character from one field to another on a game board. The goal of the game was to visit all the predetermined target fields and the distance the character moved was proportional to the strength of the BCI output. A 66% mean accuracy was achieved (Finke, Lenhardt & Ritter, 2009), which is far lower than what is expected by gamers. This result was confirmed in a later study, where the goal was to assemble a puzzle. The participants needed to select a puzzle piece that was then automatically moved to the correct position. The BCI game was played by two groups of participants (one group in single-trial mode and another in a triple-trial mode). It was found that the accuracy was 65% in the single-trial group and 81% in the triple-trial group (Ganin, Shishkin & Kaplan, 2013). These results indicate that an accuracy rate of close to 100% is not yet achievable. Interesting to note was that with the triple trial group there was a large improvement from 65% to 81%, indicating that learning took place. It can then be reasoned that although initial use of the interface did not result in very effective use, a positive increase was noticeable after several sessions. This result has repercussions when applying it to gaming, where users do not want to spend additional time on learning an interface and would rather want to use it immediately to play the game. Another limiting factor is that the participants only needed to select a puzzle

piece where-after it was automatically positioned, indicating that participants only had to activate a single command. Games generally require more than one action to be activated by the user, which could pose a problem, since the BCI has recognition of more than two mental states. Users would like the BCI to recognise a much larger variety of mental states that can be linked to actions ideally tailored for the game at hand (Lotte, 2011).

It is important that the interface should not be an obstacle to overcome before the game can be enjoyed and that the interface should allow for a range of commands to be activated during gameplay.

However, some measure of immediate control is possible, as indicated by participants making use of a BCI controlled pinball machine. Seven users with well-classifiable EEG signals were selected to participate in the study. Of the seven users, four managed to acquire good control, played very successfully and enjoyed the experience. One subject managed to get limited control and reported enjoying the game. The remaining two subjects could not establish reliable control and were excluded from further analysis (Tangermann, et al., 2009). The results of the study indicated that fast and well-timed control is possible, even though the environment is extremely rich and requires precisely timed and complex predictive behaviour.

These results may well be influenced by the limited actions that were performed by the participants as only two actions were required to operate the pinball machine. Nonetheless, they found that BCI-based pinball control is possible within the first session without the necessity to employ lengthy user training (Tangermann, et al., 2009). The study proved that BCI technology could be used effectively by a selected group of individuals without prior learning. The fact that participants with well-classifiable EEG signals were used for the study complicates the interpretation of the results. The reason that no prior learning was required during this study may be attributed to the fact that only individuals with good BCI control were used for the study. According to the results of the study, only four individuals out of seven managed to acquire good control. This suggests that only 57% of individuals with well-classifiable EEG signals could effectively use the technology. Therefore, it is unclear what the percentage for effective use would be if a random sample of participants was used, as opposed to using a sample comprising only participants who have good BCI control. Under the latter circumstance the overall percentage could potentially be far lower.

The problem with effective use for able-bodied users is further compounded by Lotte (2011) who also reported that approximately 20% of BCI users do not achieve adequate recognition performance that will enable them to control the BCI device, a problem known as BCI illiteracy. An estimated 15 - 30% of users are unable to make use of EEG-based BCIs (Blankertz, et al., 2009) and one of the greatest challenges for the adoption of EEG-based BCIs is to understand and solve the problem of BCI illiteracy (Guger, Edlinger, Harkam, Niedermayer & Pfurtscheller, 2003). It is still unclear why some BCI users display BCI illiteracy and how it can be rectified. One reason for BCI illiteracy may be the way in which the applicable part of the brain cortex is folded in relation to the scalp (Nijholt, et al., 2008) or the user's inability to imagine a certain action.

If 20% of users cannot make use of the device it will pose a practical problem as it may be abandoned quickly. This results in a challenge when trying to shift the use of BCI technology to the able-bodied user for the purpose of gaming. For an input device to be widely accepted it should allow use by any person who wishes to do so. Although successful control has been proven, the percentage of individuals who can effectively make use of the interface is limited.

User training can, however, be used to solve some types of BCI illiteracy, specifically those problems related to the user's incorrect mental execution of the task (Hwang, Kwon & Im, 2009). User training for cognitive control with BCI technology is essential, as performing a mental task to instruct a computer is new to most users. Therefore, users have to be informed of what is expected of them. Some researchers tell the user to imagine movements of their hands to focus on individual tasks, for example lifting an object. Unfortunately, this confuses a number of naive users as they are unsure of whether they should visualise the movement, feel their hand moving or if they should see someone else's hands moving. Users usually prefer the thought of moving their own hand (Plass-Oude Bos, et al., 2010). In the case of disabled users who cannot use their limbs, remnants of muscle movements can still be used to steer their thoughts (Kennedy, Bakay, Moore, Adams & Goldwaithe, 2000). Naturally, user training can be a monotonous process due to performance sometimes decreasing instead of increasing (Galán, Ferrez, Oliva, Guardia & del R Millan, 2007). Consequently it is important to keep users motivated during this process, possibly by giving feedback (Hwang, Kwon & Im, 2009).

Lotte (2011) also investigated the effective use of BCI technology in a 3D gaming environment. A number of limitations were discovered during the course of the study. Firstly, participants had to sit and not move excessively since body motions are known to generate various electrical

signals that pollute EEG signals. This poses a problem for players in terms of using the BCI for gaming purposes, as they will naturally be performing several body movements that could pollute the EEG signal and cause difficulty controlling the game characters or objects. Secondly, using BCI technology requires a lengthy calibration process for each action before the user can start with gameplay (Lotte, 2011). To work at optimal efficiency, current non-invasive BCI systems require an initial calibration process in which the system records examples of the user's EEG signals in order to tune signal parameters for that user. The BCI calibration session can take from five to 20 minutes, which may be considered too long for most computer users (Van Erp, Lotte & Tangermann, 2012). This could discourage users as they may want to use the technology immediately without the initial preparation process (Lotte, 2011). Additionally, current BCI technology is tested in unnaturally quiet environments, where the BCI still performed worse than existing game devices which include keyboards, mice and touch pads. This result may further deteriorate when the BCI is tested in real-life situations and not in a controlled laboratory environment. The technology was much slower and much more prone to errors than other devices, which makes the use of a BCI sending control commands rather inconvenient and cumbersome. Many of the limitations mentioned above, including calibration time, are currently being improved upon (Lotte, 2011).

The final result of Lotte's (2011) study was that, due to the very limited performance of an EEG-based BCI, its use as the main control device of a 3D game was rather unlikely. According to the study, there are more possibilities for BCI technology than being the main input device, for example, being an additional control channel to send commands that cannot be intuitively sent with other devices. Due to this study investigating interfaces for primary input for games, this line of investigation was not followed. Future studies may look into the possibility of cognitive input as an addition to facial recognition or other input modalities. Taking this into consideration, the widespread use of thought as a reliable method of input, especially in the world of gaming where accuracy and responsiveness is essential, does not appear to be practical as yet.

2.2.7.5. Facial expression recognition

Although necessary when using thought as input, the calibration process is not required when making use of facial expressions (Emotiv Inc, 2015). Since EMG signals, for example from eye blinking, affect EEG, most applications treat this signal as noise and attempt to detect and filter it from the data. However, the BCI allows it to be used as another form of direct input by

classifying the detected EMG into movement of facial muscle groups. The EMG from specific features is much the same between different individuals as most individuals have similar facial muscle structure. This means that training and calibration by every user is not required but can be done if higher levels of accuracy are needed (Emotiv Inc, 2015).

For gaming purposes it may be better to rather make use of an input method that does not require calibration. The selection of facial expressions as an input method for this study may also solve another problem mentioned earlier, namely the pollution of the EEG signal by body movements, as EEG was not used. Because EMG is implied to be a major source for the detection of facial features, it is also unlikely that the detection will suffer much from the BCI illiteracy problems discussed earlier (Boot, 2009).

Therefore, it can be reasoned that by making use of facial expressions instead of cognitive control the difficulties encountered in terms of BCI illiteracy, signal pollution and the need for a calibration and training process can be overcome. It can further be reasoned that using facial expressions as a method of input will be more natural as individuals make use of facial expressions on a daily basis, and this action does not have to be learned. The use of facial expressions in combination with the head-mounted accelerometer can also pose as an interface combination that can be used by disabled persons.

Although facial features may overcome some of the problems encountered with cognitive control, in order for this input method to be regarded as a possibility for gaming it has to be very effective as well. Nine voluntary facial actions were investigated using the Emotiv headset. These actions included neutral, smile, sad, surprise, angry, speak, blink as well as left and right eye movement (Heger, Putze & Schultz, 2011). Results indicated that an overall recognition accuracy of 81.8% can be achieved, which is very promising. This result was confirmed in a related study by Cernea et al. (2011), where the facial expressions of 12 participants with a basic level of computer literacy were evaluated as an accurate means of input. Results from the headset were confirmed against a more common evaluation method, in this case by using video log analysis. The participants were positioned in front of a monitor and a webcam and fitted with the Emotiv BCI headset. A random arrangement of words that represented facial expressions was displayed on a screen and the participants were asked to perform these actions. The facial actions were randomly assigned to the participants. After the tasks had been completed, the BCI data and the video logs were compared. The BCI headset results showed that the correct detection of the facial expressions varied between 70-100%,

depending on the particular facial expression (Cernea, et al., 2011). This shows that facial recognition can be used on demand and has a high recognition accuracy. Further research should be conducted based on the successful results from the study by Heger et al. (2011) and Cernea, et al. (2011), focusing on how well these results will translate in the use of facial expressions to accomplish a task.

Whether the same percentage of success will be achieved when applying facial recognition to control applications and not only testing the recognition rate must still be investigated. It is important to note that when using facial expressions to control a device or application an additional layer of effort is added, essentially the users' attention is divided between the application and the BCI itself (Leeb, et al., 2013). Thus, it would be interesting to investigate the use of facial expressions during gameplay where the user will have to concentrate on the use of the interface as well as the game itself. This added difficulty may have a negative influence on the effectiveness of the interface. This study addressed the above shortcoming by including game related tasks during user testing. As a result, the usability of facial expressions as a reliable method of input was investigated under simulated gaming conditions.

The difference between the use of cognitive and facial expression control to complete tasks and not just the recognition thereof is clearly illustrated in a study by Kester (2012), where participants were divided into two groups, one group using facial expressions and the other group using cognitive control to complete a set of point and click tasks. The participants in the facial expressions group were able to complete every task, while the participants who made use of the cognitive function of the Emotiv BCI were unable to complete any of these tasks (Kester, 2012).

Therefore, facial expressions as a control technique seems more promising for further investigation than the use of thought, in particular within a gaming environment, as accurate and reliable control is an essential requirement. Gamers require near perfect accuracy in order to be competitive in the specific game that they are playing. Gamers expect the interface to respond every time they issue an instruction, and this does not seem to be the case when utilising cognitive BCI input.

There are, however, factors to be considered when making use of facial expressions for input purposes. For instance, in a study by Boot (2009), data was recorded using 32 EEG electrodes and 8 EMG electrodes while 10 participants performed 4 different facial expressions. Results

indicated that facial expressions could be classified from data recorded by EEG sensors with an average classification accuracy of over 80%. However, facial expressions with overlapping muscles led to lower accuracy levels. When these overlapping facial expressions were avoided the accuracy improved to over 90% (Boot, 2009).

Using a BCI to circumvent direct real world interaction may also seem unnatural to most users. Users are accustomed to converting their thoughts into bodily actions but not directly manipulating the environment around them with their thoughts. This means that a user will need to develop new levels of control when using brain activity directly (Plass-Oude Bos, et al., 2010). This is in contrast to what NUIs are defined as, namely that the action has to be natural to the user (Steinberg, 2012). Everybody naturally uses facial expressions and a person does not have to learn it before he/she can use it as input. Thus, facial expressions, which are used by individuals on a daily basis, may prove to be an effective primary input method in the field of gaming and needs to be investigated further. Therefore, command activation can be achieved to some extent with both cognitive control and facial expression recognition. Nonetheless, command activation only forms part of game input since cursor control is also a very important factor. For the purposes of this study, the Emotiv's facial expression functionality for command activation was combined with an alternative method of cursor control.

2.2.7.6. Alternative cursor control

With the BCI's cognitive and expressive capabilities replacing the keyboard, a natural alternative to the mouse is also required. In an attempt to replace the mouse with an electrode headset similar to the Emotiv BCI, several objects were displayed around the edge of a computer screen, although only one of the objects was chosen as the target to where the cursor had to be moved. The task that was assigned to the participants was to move the cursor from the centre of the screen to the correct object by making use of EEG. An additional EEG feature was then used to select the object. The goal of this task was to emulate mouse operation by making use of EEG signals via a BCI headset (McFarland, Krusienski, Sarnacki & Wolpaw, 2008), which is similar to the multi-directional tapping test based on ISO 92 41-9 (ISO 9241, 2000). Although the participants completed between 14–38 training sessions during the study, the data generated during the last three sessions were used for statistical analysis. In 59–88% of the trials the participants reached the indicated target and in 71–91% of the trials the

participants correctly selected or rejected the target (McFarland, Krusienski, Sarnacki & Wolpaw, 2008).

The results showed that participants could learn how to use scalp-recorded EEG signals in order to move a cursor towards a target object and then also select the target. The participants gradually, over several training sessions, acquired better EEG control. It was also observed that as the levels of control improved, the motor imagery that users utilised during the early training sessions (for example, the accompanying movement of the hands) became less prominent (McFarland, Krusienski, Sarnacki & Wolpaw, 2008). This indicates that participants developed better cognitive control, thus eliminating any muscle movement and relying only on their thoughts. If individuals are able to make use of an EEG system without any accompanying muscle movement the possibility of using additional input mechanisms in conjunction with the EEG device, for example using the mouse, keyboard and EEG device concurrently, becomes a viable option.

Thus, according to this study EEG control of a mouse cursor is possible, although there is a significant training process attached to these results. For the purposes of the current study, which falls in the field of gaming, these results would not facilitate effective control. The long training period, combined with the better than chance level of control acquired, would not make the technology feasible for gaming purposes where very accurate and fast control is not negotiable.

More promising results were obtained with a hands-free mouse emulator using the Emotiv EEG headset. The mouse pointer was moved in response to head movements as obtained from the accelerometer and mouse clicks were activated by the user blinking. The results showed that mouse control was successful when the accelerometer was included in the headset. Despite movements from the participants, the study showed excellent results during double blink events, with an average detection rate of 94.9% (Rosas-Cholula, et al., 2013).

Thus, the head-mounted accelerometer can be successfully used for cursor control. It has also been indicated that blinking could be successfully used to activate a command. During this study, the head-mounted accelerometer was combined with command activation through facial expressions, as well as glove-based input (Section 2.3). By making use of the head for cursor control the user was able to use common actions, such as turning the head left and right, or

moving it up and down, in order to control the cursor – which in this instance is aiming or looking at a target. These actions may be more natural and intuitive than using a mouse.

The Emotiv BCI mouse emulator has different sensitivity settings that can be used when controlling the mouse. These settings range from low to high sensitivity, with low sensitivity requiring larger movement of the head than high sensitivity in order to move the cursor a set distance. During the pilot study (Chapter 5), it was determined which of the sensitivity settings was the most effective to use during the user testing. Low, medium and high sensitivity options was investigated.

2.2.8. Summary of BCI technology

Facial recognition proved to be a more accurate option than cognitive input for use as a primary input mechanism in the world of gaming. EEG-based input still has the potential to be used as an additional input mechanism for non-direct interaction by the user, for example, detecting the user's mental and emotional state and using this information to alter certain aspects of the game in accordance with the change in emotional status. This may improve the immersion the user experiences during gameplay and will not affect the input accuracy. Only primary interaction methods were investigated for the purpose of this study. Therefore, the Emotiv BCI's Expressiv suite was investigated instead of the Cognitiv suite which is better suited to being an additional input channel. During this study, the facial recognition of the BCI as well as its head-mounted accelerometer to detect head movement was used.

This study focussed on BCI use for able-bodied individuals in a gaming environment. Although the study used methods and techniques from the field of BCI, it is not purely a BCI project. Where BCI research mainly focuses on interpreting a user's thoughts, this study focussed on signals from the facial muscles.

Glove-based input technology pertaining to the study at hand will be discussed in the following section.

2.3. Glove-based technologies

2.3.1. Introduction

We use our hands to interact with and manipulate our environment on a daily basis, especially when manipulating a digital environment (Kumar, Verma & Prasad, 2012). Currently, traditional human-computer interface devices constrain the user's interaction with computers and computer applications. In an effort to change this, detection directly from the user's hands is being investigated, freeing the user from the constraints of traditional devices such as keyboards, mice and joysticks (Sturman & Zeltzer, 1994). Direct input from the user's hand can be derived from either hand postures or hand gestures (LaViola, 1999). This type of input may be more intuitive and natural for users.

2.3.2. Hand postures and gestures

While hand postures and gestures are usually believed to be the same, certain characteristics separate these actions. A hand posture can be defined as a static action. An example would be to make a fist and holding it in that position. Postures can be categorised as simple or complex postures. When performing a simple posture, each of the fingers are either fully extended or flexed. When the fingers are bent at any angle, which does not include fully extending or flexing them, the action is referred to as a complex posture (Murakami & Hitomi, 1991).

Conversely, a gesture can be defined as a dynamic action, which can be either simple or complex. Executing a simple gesture can be accomplished by performing two different types of hand movements. The first method includes actions such as performing a pinching posture and then changing the hand's position in space. The second method requires the fingers to be moved with no adjustment of the position or orientation of the hand. A complex gesture can be defined as an action that is comprised of finger and wrist movement, as well as adjusting the position and orientation of the hand (LaViola, 1999).

Hand postures and gesture recognition can be classified into two categories, namely the camera vision method (Rekha, Bhattacharya & Majumder, 2011) and the data-glove method (LaViola, 1999). With the camera vision-based method a camera is used to detect and recognise different hand commands (Rekha, Bhattacharya & Majumder, 2011). The data-glove method makes use of wired gloves to recognise different hand-based commands (LaViola, 1999). The data-glove

based gesture recognition method can be successfully implemented in games due to its high accuracy compared to the camera vision-based methods (Heo, Lee, Park, Kim & Whang, 2010). Therefore, during this study, the focus was on the data-glove method.

2.3.3. Glove-based devices

A wired glove, also referred to as a data-glove or cyber-glove, is an input device for HCI worn like a glove. Sensors are utilised to capture data, for example, the bending of the user's fingers. Motion trackers, magnetic trackers or inertial tracking devices can be attached in order to detect the position and rotation of the user's hand. Certain gloves can offer haptic feedback that will simulate the sensation of touch and can be used as a form of output (Adnan, et al., 2012).

A glove-based system is defined as a system that is comprised of an array of sensors, electronics for data acquisition and processing, and a power supply, all combined within a glove worn on the user's hand. Typically, the glove is made of cloth and records data related to the user's hand motion while being worn (Dipietro, Sabatini & Dario, 2008). In order to detect user intentions, different recognition algorithms are applied to analyse the raw data, which will be used to instruct the device in question (LaViola, 1999).

Glove-based input devices allow us to apply our manual dexterity to the task that we are performing (Sturman & Zeltzer, 1994). This leads to the idea that glove input may be more natural and intuitive than existing interfaces, due to the use of an individual's hands as a primary source of input. A substantial amount of research has been dedicated to developing technologies for supplementing our hand-coordinated abilities. Glove-based technology started in 1970 when the first glove system was designed. Since then, there have been different glove designs proposed over the years, with some of these designs being better suited to certain applications than others (Dipietro, Sabatini & Dario, 2008). Various applications of glove technologies have been investigated, including projects aimed at natural interfaces, systems for understanding sign languages as well as robotic control (Sturman & Zeltzer, 1994). Some of these applications will be discussed in the following section.

2.3.4 Applications for glove-based devices

2.3.4.1. Designing

Traditionally glove-based devices are used during interaction with virtual environments. Making use of a computer screen the user can envision environments that are being designed before the actual construction thereof, eliminating the requirement of expensive prototypes (Li, Lau & Ng, 2003). The design process is made more natural by utilising gloves to directly create 3D shapes (Keefe, Feliz, Moscovich, Laidlaw & LaViola, 2001). This is done by grasping virtual objects or by using gestures to interact with the computer (LaViola, 1999). Thus, by making use of the glove-based device users can utilise a 3D space to produce a device, where-after the user could evaluate the mechanism's motion by moving around it (Evans, Vance & Dark, 1999).

Classic uses of the device include pointing with the index finger in order to change the users' view in the virtual environment (Krapichler, Haubner, Loesch & Englmeier, 1997), or the use of the pinching gesture to select a menu option (Piekarski & Thomas, 2001). Among the first in the industrial arena to utilise glove-based technology for simulated environments was Daimler-Benz and Boeing. By utilising data gloves, employees at Daimler-Benz could choose between diverse furnishing selections for Mercedes interiors. Boeing made use of glove-based technology to simulate maintenance tasks such as loading weapons or removing parts that needed to be repaired on an aircraft (Steffan, Schull & Kuhlen, 1998). Glove-based technology has also been used in virtual training environments for pilots (Doerr, Rademacher, Huesgen & Kubbat, 2007), soldiers (Xu, Yao & Zhang, 2006) and astronauts (Ronkko, et al., 2006).

2.3.4.2. Information visualisation

Computer graphics are regularly utilised to generate visual depictions of data to assist in the interpretation thereof. This technique is especially useful when making use of complex numerical representations of data (Durlach & Mavor, 1995). By utilising glove-based technologies, the user's interaction with the data will be more natural. This can lead to the improvement of traditional data visualisation methods (Dipietro, Sabatini & Dario, 2008).

This technique was used by NASA in the late 1980s in a virtual wind tunnel, where the users could visualise a computer-generated airflow around an aircraft. By utilising a data glove, the

users could move and observe the changes in one or more streamlines of the fluid flow. The results of these changes in flow were computed and visualised in real time by supercomputers (Bryson & Levit, 1992). Data visualisation systems have also been used to manipulate geospatial (Shahabi, et al., 2006) and medical data (Tani, Maia & von Wangenheim, 2007), as well as general computer files (Deller, Ebert, Bender & Hagen, 2006).

2.3.4.3. Robotics

Robot programming could be made easier and more natural by using glove-based systems in relation to automatic programming. This is very useful when utilising systems where the control of a large number of joints is required, which is usually very difficult to achieve when making use of standard control methods. The robot learns certain movements automatically by detecting a demonstration of the action by an individual (Biggs & MacDonald, 2003). A number of researchers have used glove-based technology to control multi-fingered robotic hands (Pao & Speeter, 1989). An impressive application was implemented by NASA and the Defence Advanced Research Projects Agency (DARPA), where the Robonaut, a human scale robot, was developed to assist in extravehicular activities - a dangerous activity usually performed by astronauts. This allowed the user to instruct the robot to execute certain maintenance tasks, for example, threading nuts into bolts or tying knots (Diftler, Culbert, Platt & Bluethmann, 2003).

2.3.4.4. Sign language

Utilising hand posture and gesture recognition to detect sign language are natural applications of this technology (LaViola, 1999). Several glove-based devices have been developed for the automatic recognition of sign languages used by hearing impaired individuals (Sturman & Zeltzer, 1994). These devices varied in features, for example, the number of recognisable signs, types of signs and the percentage of signs correctly interpreted (Dipietro, Sabatini & Dario, 2008). Sign language could be translated using embedded interfaces that convert sign language into text or vocal outputs (Gao, et al., 2000), for example, the Talking Glove recorded, recognised, and translated American Sign Language into text or spoken words by using a CyberGlove (Unites States of America Patent No. 5,047,952, 1989). Converting gestures into

speech could assist hearing impaired individuals in communicating with individuals who are not versed in sign language (LaViola, 1999).

2.3.4.5. Medical applications

Glove-based systems were not initially intended for use in the medical field, nonetheless they have been attracting increased attention from researchers in the field. Initial research focused on automatic sign language recognition, however, researchers have been exploring several different avenues including motor rehabilitation, the analysis of motor rehabilitation, ergonomics and medical education and training (Dipietro, Sabatini & Dario, 2008).

2.3.4.6. Entertainment

Glove-based input technology have been utilised during the production of movies (Minoh, Obara, Funtatomi, Toyura & Kakusho, 2007), as well as in gameplay and for the animation of virtual characters (Adamo-Villani & Wilbur, 2007). Glove technologies have also been applied to control acoustic levels during musical performances (Morita, Hashimoto & Ohteru, 1991). A system was developed where two CyberGloves were used to manipulate 3D computer-generated objects. The position, orientation as well as the shape of the virtual objects could be adjusted using hand gestures, where-after the alterations to the virtual objects were mapped to changes in acoustic levels (Mulder, Fels & Mase, 1997).

Glove-based interaction has also been applied to control audio and video devices. An interaction system, where a television was operated by making use of hand gestures, was developed by Freeman and Weissman (Freeman & Weissman, 1994).

2.3.4.7. Wearable and portable computers

One of the latest advances has been the introduction of glove-based technology as interaction devices for end user electronics, for example, for use as text input and pointing devices for portable and wearable computers. While exploring the possibilities for more convenient interaction devices for portable and wearable computers, developers initially focused on reducing the size of traditional input devices. This approach had its limitations as these devices cannot be reduced beyond a certain point. These devices still need to be big enough to be easily seen, and operated using the user’s hands. After these setbacks the developers focused on a new approach, where they aimed to replace traditional input devices with wearable technology.

Glove-based technology emerged as a likely solution to achieving this goal as gloves allowed hand postures to be captured and then converted into keystrokes and commands. The necessity for a traditional keyboard could then be excluded (Dipietro, Sabatini & Dario, 2008). Thus, it seems very promising to use a glove-based input device to replace the role of the keyboard as a primary method of command activation during gameplay. Therefore, a data glove in combination with the head-mounted cursor control that the Emotiv BCI offers can be used as a more natural replacement for the keyboard and mouse. Making use of these two methods of input will result in a new multimodal input device.

2.3.4.8. Multimodal interaction

Multimodal interfaces is an area in which glove-based technologies can play an important role. Hand posture and gesture detection used in combination with speech have been applied in several systems in order to create a natural multimodal interface. Using more than one modality, results in the one complimenting the other and eliminating specific detection drawbacks of a certain modality (Oviatt & Van Gent, 1996).

This study focussed on comparing multimodal NUIs to the traditional keyboard and mouse combination in terms of gameplay. The glove-based device was used in conjunction with a head-mounted accelerometer to form a natural multimodal interface. Although various glove designs were developed over the years, only a few were made commercially available (Dipietro, Sabatini & Dario, 2008). These glove designs will be discussed in the following section.

2.3.4.9. Gaming applications

There has been longstanding use of glove-based input technology in the entertainment industry (Dipietro, Sabatini & Dario, 2008). Although glove-based input technology has been previously used for the animation of virtual characters (Adamo-Villani & Wilbur, 2007), according to Orozco, et al. (2012) recent research has shifted more towards haptic research in the area of home entertainment, especially computer games. A number of gaming focused gloves (see section 2.3.5 for a discussion on some of these) have been developed, however, as far as can be ascertained, there have been few recent formal research studies where a glove has been specifically used in the computer gaming field.

The Peregrine glove has been informally tested in League of Legends and DotA2, which are very popular online games. Over 1,000,000 actions were tested by 5 different gamers and showed an action per minute increase of 5 - 20% over traditional interfaces such as a keyboard and mouse. If properly used, the glove will result in greater focus, increased in-game awareness, as well as quicker reaction times when activating abilities and items (Iron Will Innovations, 2013). These improvements may lead to more immersive and natural gaming, which motivates the need to test the use of glove-based technology specifically for gaming, in order to confirm these positive, informal findings.

2.3.5. Commercially available glove technology

The first glove for computer input was developed in 1977. Many others were developed and tested thereafter (LaViola, 1999), including the AcceleGlove (Hernandez-Rebollar, Kyriakopoulos & Lindeman, 2002), the LightGlove (Howard & Howard, 2001) and the StrinGlove (Kuroda, Tabata, Goto, Ikuta & Murakami, 2004). The most recent commercially available glove-based technologies will be discussed in this section. This study focussed primarily on glove-based input devices that were designed and marketed as a gaming input device.

CyberGlove

Developed by James Kramer at Stanford University and commercialised in 1992, the glove is equipped with 18 piezo-resistive sensors. This device is comprised of two bend sensors on each finger and four abduction/adduction sensors. It also measures thumb crossover, palm arch, wrist flexion and wrist abduction/adduction. Software called VirtualHand is used for the calibration process in order to customise the device to the differences between user's hands, finger length and thickness. A wireless CyberGlove as seen in Figure 2.3.1, called the CyberGlove II, was commercially released in 2005. The CyberGlove was considered to be one of the most accurate glove systems available at the time (LaViola, 1999). The CyberGlove and CyberGlove II have been successfully utilised during several studies (cf. Adamovich, et al., 2004; Schabowsky,

Figure 2.3.1 - CyberGlove (Sesto Senzo, 2012)

Godfrey, Holley & Lum, 2010). This device was not included in the study as it was not available to the researcher.

DG5 VHand 3.0

The DG5 VHand 3.0 data glove (as seen in Figure 2.3.2) has five embedded bend sensors that it utilises to accurately measure finger movements. The glove also has embedded motion sensors that allow it to detect both hand movement and hand orientation. It communicates with external devices through a USB cable or a WiFi connection. The device can be powered through the USB connection or use its internal battery for complete wireless use (Virtual Realities Ltd, 2013).

Figure 2.3.2 - DG5 Vhand glove 3.0 (Virtual Realities Ltd, 2013)

This input device has been investigated numerous times with good results (cf. Camastra & De Felice, 2013; Kumar, et al., 2012). However, it was not included in the current study as it was not accessible to the researcher at the time.

The Peregrine gaming glove

According to the manufacturer, the Peregrine gaming glove has 18 touch points and 3 activator pads that allow for more than 30 user-programmable actions. The Peregrine glove, as seen in Figure 2.3.3, has been developed and is marketed as a gaming glove. By simply touching the user's fingertips with his thumb a user can access 30 programmable instructions or characters (Peregrine, 2013). This feature makes the glove an ideal input device for gamers as they can customise the glove, and may thus be able to activate commands faster than with a keyboard. This will make it easier for users to control their game characters and allow for easy and fluid transition between different commands, which will therefore lead to the user having a competitive advantage.

Figure 2.3.3 - The Peregrine gaming glove (Peregrine, 2013)

Whenever these activator pads make contact with a touch point on a finger, a keystroke is instantly sent to the application in use. This glove does not detect finger flexure, instead it makes use of certain points of the glove being touched or activated by the activation pads. An advantage of this glove design is its low cost, thus overcoming a big hurdle to everyday use of a device, namely its affordability (Peregrine, 2013).

In order for glove input to be investigated during this study it has to be considered a natural input technique. The data glove was initially developed to act as a method of input during gameplay and it was concluded that the glove enabled a more intuitive, as well as a more immersive experience for the user. Furthermore, the potential applications of glove-based technologies are widespread and not just limited to being and input device for games (Foottit, Brown, Marks & Connor, 2014). This study confirms that glove-based input is more natural and intuitive than using traditional interfaces and thus further supports the use of glove input as part of a natural, multimodal interface.

Two wireless CyberGloves were utilised in a study by Logsdon (2011) to track the hand postures of participants for gesture recognition and two trackers were used for tracking wrist motion. A 3D-environment was developed using the Microsoft XNA game development platform. Each participant had video game experience using a variety of controller types, including keyboard and mouse, game pads, Wii Remotes and Kinect (Logsdon, 2011).

Among the tasks that were performed were two that have importance to the study at hand: the participants were instructed to choose and select certain menu items, as well as play a first person shooter game. Data was recorded for both stationary and moving targets during the selected game (Logsdon, 2011). Due to stationary and moving targets forming a large part of most games, it would be important to select appropriate tasks that include both types of targets in order to compare the NUI with the traditional keyboard and mouse in the current investigation.

The participants found it intuitive to select the menu items by making use of a grab gesture with their hand. The only difficulty that was experienced was that participants occasionally selected the menu item above or below the expected menu item. This lapse in accuracy was caused by slight hand movement triggered by the grab gesture. This may be caused by using the same modality for command activation and cursor control. Thus, it may be reasoned that positioning the cursor as well as activating the select command with the gloves caused these

instability issues. This is an important finding to take into consideration for the current study: it may be more feasible to rather use another modality for cursor control and use the glove only for command activation. Therefore, for the purpose of this study the glove was combined with the head-mounted accelerometer, thus combining to form a multimodal NUI. This may lead to the elimination of the accuracy problem encountered during the study by Logsdon (2011).

When playing the first person shooter game, participants indicated that the method utilised for activating the shoot command was very natural and intuitive due to the "trigger" gesture employed with the glove being very similar to real-life shooting (Logsdon, 2011). This is also a very important finding in terms of the study at hand: during this study a 2D shooter game will be used, thus, when using the glove as input it will be advantageous to also make use of an action that is similar to the real life act of pulling a trigger.

Participants indicated that they enjoyed finger-based interaction more compared to current methods of video game interaction that include both gesture and non-gesture based input methods. They also thought that this method of interaction would be very good for certain games, for example shooting games (Logsdon, 2011). It would be interesting to establish whether these positive findings concerning user satisfaction will be repeated during the current study, which will also measure user satisfaction and is essentially a shooting game.

The conclusion made was that by making use of finger interaction, the user's hands can now be utilised as an input method for games where high performance and precision is required. It was also argued that much more could be gained by researching this area of interaction (Logsdon, 2011). Therefore, as mentioned earlier, a data glove was used during the current study to attempt to replicate these positive results. However, there was a different modality used for cursor control as in this case, the glove was combined with a head-mounted accelerometer. This will hopefully solve the difficulty experienced with aiming accuracy that was established by Logsdon (2011) during his study, when both command activation and cursor control was conducted using a glove.

In a recent study, a glove that operated on exactly the same principles as the Peregrine glove was developed and tested. This glove also makes use of electrical contacts between the user's fingers, and was investigated in order to test the viability of replacing a keyboard with a one-handed glove for input. Participants were asked to type a short sentence by making use of the

glove. The time to complete the task as well as the number of errors were measured and used for statistical analysis (Peshock, Dunne & Duvall, 2014).

The results of the investigation indicated that after each session the input speed increased while the errors decreased, indicating that a learning curve was present. During the study it was concluded that the glove has the potential to be implemented in various fields as it could successfully be used to replace keyboard input (Peshock, Dunne & Duvall, 2014). This indicates that in a gaming environment a similar glove can be used to replace the keyboard. Consequently, the Peregrine glove posed to be a suitable, natural alternative to the keyboard, and an appropriate choice as it is also marketed as an input device specifically focused on gaming (Peregrine, 2013).

The importance of adding feedback to the interaction was also emphasised, as it would help improve accuracy as well as user satisfaction. The Peregrine glove does not have any feedback functionality. This was added to the application. During this study sound was used as a means of feedback as each action triggered an accompanying sound. This gave the participant feedback on whether or not the command was successfully activated.

2.3.6. Summary of glove-based input

Taking the abovementioned information into consideration, the Peregrine gaming glove was selected as it is a recent, commercially available gloved-based input device that is specifically focused on gaming, can be regarded as a natural interface and is available to the researcher.

The Peregrine gaming glove in conjunction with the head-mounted BCI accelerometer was evaluated as a combination during gameplay. The results were compared to the results of the traditional human-computer interface in order to explain the continued use of traditional user interfaces over NUIs such as the Peregrine glove.

2.4. Conclusion

This chapter discussed HCI and NUIs which included BCIs and glove-based input. After investigating the different input devices that are currently commercially available, two have been selected to be included in the study at hand. The Emotiv BCI was utilised, focusing on its facial recognition function as well as its mouse emulator. The second NUI to be included was the Peregrine gaming glove.

The first multimodal interface was the combination of facial expressions for command activation together with the head controlled mouse for cursor control. The second multimodal interface included the Peregrine glove, for command activation, in conjunction with the BCI's head controlled mouse. These two NUIs were compared to the traditional user interface consisting of the keyboard and mouse in order to answer the research questions posed in Chapter 1.

The following chapter will include an in depth discussion on the research design that will be followed in this investigation.

CHAPTER 3 - Research design

3.1. Introduction

The preceding chapter discussed some of the available literature which was used to motivate this study. This chapter will discuss the experimental research design which was utilised to attempt to provide answers to the research questions posed. Details concerning the usability evaluation methods that were conducted and the procedures that were followed, will be discussed during this chapter.

3.2. Research problem

As mentioned in Chapter 1, although a large percentage of game console users have started using NUIs for gaming purposes (Oikonomidis, Kyriazis & Argyros, 2011), little change has been perceived in the personal computer (PC) gaming scene in terms of NUI use (Beckhaus, et al., 2005; Kumar, 2014). Most PC gamers are still using the traditional keyboard and mouse input combination, which have been known to cause repetitive strain injuries (Ramos, James & Bear-Lehman, 2005). No clear reason why NUIs have not migrated to PC gaming could be found in the literature studied. The problem might lie in the usability of these interfaces within a PC gaming context. This study thus proposed to investigate the usability of NUIs in a gaming environment.

3.3. Research question

As mentioned in Chapter 1, in order to investigate this problem the following research question was posed: To what extent is the usability of a game influenced by the use of a NUI as opposed to a traditional keyboard and mouse combination? Therefore, this research question is causal in nature. Since experimental research can be used to determine whether a cause-effect connection is present in one element or between a set of elements (Shadish, Cook & Campbell, 2002), in this case between the different interfaces, an experimental research design was used in this study.

3.4. Experimental design

As explained by Mouton (2001) the experimental research design can be described as a study that is usually quantitative in nature and encompasses a causal study that is undertaken under highly controlled conditions. A hypothesis first has to be established before an experiment can

be conducted and through the experiment a cause and effect relationship can then be established (Faulkner, 1998). The hypothesis for the current study stated: There is no difference between the usability of a NUI and GUI when used for gaming purposes.

An experiment can be classified as the gathering of data by using the population that is being studied or controlled during the investigation. The data is usually in the form of measurements and observations (Dowdy & Weardon, 1983). An experimental design usually has the highest internal validity, which guarantees that what happened was the cause of what had been observed (Trochim, 2005). External validity is the degree to which the conclusions of the study can be generalised to other people, settings and times (Calder, Phillips & Tybout, 1982). By conducting the experiment in a controlled setting the researcher maximises internal validity at the expense of the external validity. Nonetheless, an experiment is still considered the most robust and best research design when conducting a quantitative study (Trochim, 2005).

There are two types of data that can be used to answer research questions, namely primary and secondary data. Primary data can be defined as original data that has not been published before. Primary data can be obtained through experimentation, observation or surveys (Mouton, 2001). Secondary data can be defined as data that was collected by an individual other than the researcher, and not collected for the specific research at hand (Frankfort-Nachmias & Nachmias, 1992). During this study data specific to this study was gathered, and was thus primary data. Taking this into consideration, it is important to note that the experimental research design makes use of empirical data that can be collected as primary data (Mouton, 2001). Data that is collected using observations and measurements of reality is referred to as empirical data (Trochim, 2005). The data for this study is empirical since the data was gathered using user testing and questionnaires.

Making use of controlled experiments as an approach to HCI research has been adopted from research methods in psychology and are commonly used to evaluate interfaces (McGuffin & Balakrishnan, 2005), users' styles of interaction (Moyle & Cockburn, 2005) and to understand the thought process in the framework of system interaction (Li, Cox, Blandford, Cairns & Abeles, 2006). Therefore, this design is suited to the current study which falls in the discipline of HCI.

Experiments in HCI are largely considered to involve usability testing and have been found to be particularly useful and applicable to HCI research (Faulkner, 1998). When experts are

involved in usability testing the method is known as an expert review, therefore not making use of users when conducting testing (Molich & Jeffries, 2003). When models are used during usability testing the process is known as model-based evaluation (Utting & Legeard, 2010), whereas user testing (Nielsen, 1992) make use of users (instead of experts or models) to test a user interface.

In order to compare the usability of the different interfaces there had to be a clear understanding of what usability actually refers to, how usability is defined, as well as the usability model that was used during this study. These details will be discussed in sections 3.5 and 3.5.1.

3.5. Usability Testing

Experiments in HCI can take on the form of comparative experiments, where similar existing systems are tested and compared to each other. It can also take on the form of an absolute experiment, where the system is tested on its own (Faulkner, 1998). As stated previously, the experimental design was implemented during this study, more specifically, user testing was performed in order to compare the different interfaces in terms of usability.

Measuring user satisfaction is an important factor before introducing market ready goods and services. Due to this, the evaluation of products has been an essential step for development in the field of HCI (Cernea, Olech, Ebert & Kerren, 2011). The focus of product development must be on the users of the product – the developers may be too far removed from the actual day to day tasks that will be performed by the users and they may not see defects in the product that will be clearly visible to the users. It is very important to note that users and not managers or developers decide whether a product is usable (Dumas & Redish, 1999). Thus, usability testing was conducted in order to assess the users' satisfaction while making use of a product. However, usability testing is much more than determining satisfaction, as it includes investigating the efficiency, effectiveness as well as the learnability of a product (as mentioned earlier).

During this study participants, representative of end-users, took part in the evaluation of the selected interfaces. This study thus made use of user testing. The experimental research design was applicable since this study aimed to compare the usability of two natural, multimodal interfaces with that of the traditional keyboard and mouse combination.

User testing is a common tool used to evaluate the usability of a product where users are given tasks in a test environment and encouraged to think aloud while trying to accomplish the tasks. By doing this, information is gathered on how the user interface matches the natural human way of thinking and acting, and exposes the features of the product that need to be improved (Rubin & Chisnell, 2008). It is essential that the most important tasks are requested to be completed first, due to participants not always finishing all the tasks assigned to them (Dumas & Redish, 1999).

The administrator facilitates the usability testing sessions using a test script, while also taking detailed notes of the participants' behaviour and comments. At the end of the testing session the participants either complete a questionnaire or have a debriefing interview about their experiences and opinions in terms of the product (Rubin & Chisnell, 2008). The entire process usually takes four to six weeks, including reporting the results. Usability tests are usually conducted in usability test laboratories, consisting of a simulated testing area connected to a monitoring area with a one-way mirror. During usability testing, participant screening questionnaires and recruiting scripts can also be used (Kaikkonen, Kekäläinen, Canka, Kallio & Kankainen, 2005).

When conducting usability studies, the gathering of the data can be conducted under controlled conditions known as laboratory-based usability testing or under uncontrolled conditions which is known as field-based usability testing. The complications in conducting usability testing and collecting data are greatly reduced when performing the test in a controlled laboratory compared to field-based usability tests (Kjeldskov & Stage, 2004). Field-based usability testing is conducted in an environment where conditions are unpredictable since the tests are conducted outside of the laboratory, in a real life environment. For example, noise and lighting conditions may change as the usability test progresses and this may influence the results (Kjeldskov & Stage, 2004). For the purpose of this study laboratory-based usability testing was utilised, due to the need for participants to be tested one by one in a controlled environment.

Working one at a time, the test participants all performed the same tasks under controlled conditions. These tasks were representative of those typically performed on the system and can be compared to the specific goals that have to be met as stated in the definition for usability. A test plan, referred to as a protocol, usually describes the complete user test.

In this study three different user interface combinations were compared to each other in terms of usability. As measurable data was needed for this comparison, data was captured through user testing that was conducted on the different user interfaces.

In order to make effective use of usability testing, the definition of usability has to be clearly understood and will be discussed in the next section.

3.5.1. Usability

Users use products to be more productive, in other words, less time should be spent on a task when making use of a specific product versus doing it manually or by using a similar product. The hardware and software that the users use are thus tools to achieve their goals in the shortest time possible (Dumas & Redish, 1999). The way in which users interact with devices is important because the risk does exist that users will become impatient with an unintuitive, unnatural interface and will simply not use it anymore. Because of this problem developers have to design user interfaces that are more natural and easier to use (Steinberg, 2012).

According to the International Standards Organisation (ISO), usability can be defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO 9241, 2000). This section of ISO 9241 introduces the concept of usability but does not make specific recommendations in terms of product attributes. According to Shackel (1991), usability can be described as "the capability in human functional terms to be used easily and effectively by the user, while giving the user specific training and support, to fulfil the specified range of tasks, within the specified environmental scenarios" (Shackel, 1991, pg. 24). Other usability definitions include those of Constantine and Lockwood (1999), Shneiderman (1992) and Preece et al. (1994).

The definition of usability that was used during this study was a combination of the definitions listed above. In this study, usability referred to the extent that a product can be used by a specific user to achieve a specified goal with effectiveness, efficiency and satisfaction, while giving the user specific training and support. Four attributes were taken from the consolidated definition for a product to be declared usable, namely effectiveness, efficiency, training (learnability) and the users' satisfaction level. This closely resembles the quality components

proposed by Nielsen (Nielsen, 1994 as cited by Abran, Khelifi, Suryn & Seffah, 2003), as seen in Figure 3.1.

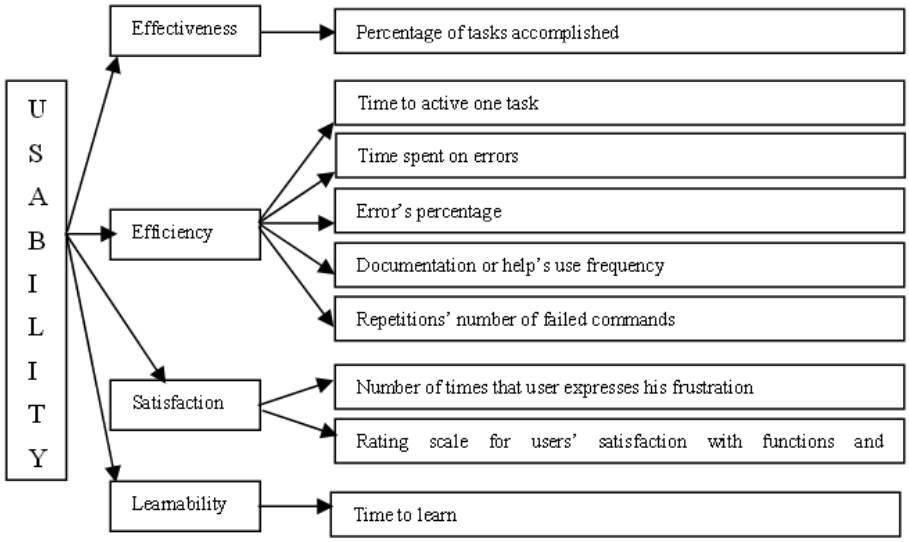

Figure 3.1 - Nielsen's Usability Model (Abran, et al., 2003)

There are three different aspects of user activity during the research and design of HCI. These aspects are categorised as the physical, cognitive and affective facets of HCI. The physical aspect determines the technicalities of the interaction between the user and the computer. The cognitive aspect takes into consideration how the user understands and interacts with the system. The affective aspect of HCI is a more recent development and its goal is to make the interaction a pleasurable experience for the user. The affective aspect of HCI also has the objective of affecting the attitudes and emotions of the user (Karray, et al., 2008).

Therefore, in order to fully understand the usability of a user interface, the physical, cognitive and the affective aspects of the interaction needed to be investigated. Nielsen's usability components (Figure 3.1) were matched to these aspects as follows:

- When focusing on the physical facets of usability the researcher could investigate the effectiveness as well as the efficiency of the interface.
- When focusing on the cognitive facets of usability the researcher could investigate the learnability and memorability of the user interface.
- When focusing on the affective facets of usability the researcher could investigate the satisfaction of participants with the interface.

Nielsen's model was consequently utilised due to it addressing all aspects of usability as was previously defined in this section. It can also be closely matched to the three different aspects

of user activities. Data regarding effectiveness, efficiency, satisfaction and learnability were collected while testing the different user interfaces. The data was then utilised to compare the different interface combinations. In order to gather data for the traditional keyboard and mouse combination, as well as the two NUIs, user testing had to be conducted.

Nielsen's model also incorporates memorability as a component of usability. However, memorability was not evaluated during this study due to time constraints. The usability components according to Nielsen's usability model (Nielsen, 1994) was investigated as follows:

Effectiveness

Effectiveness is how well the user is able to achieve the required task by using the interface (ISO 9241, 2000), and can be measured in terms of accuracy and completeness (Cato, 2001). During this study effectiveness was measured by taking into consideration the number of tasks successfully accomplished (which correlates with the "Percentage of tasks accomplished" as seen in Figure 3.1).

Efficiency

Efficiency is the amount of resources that is required to complete the specific task (ISO 9241, 2000). These resources include measurements such as time, money or mental effort (Bevan & Macleod, 1994). It is typically investigated by determining how quickly a user can perform a task after having learnt to operate the device or application (Nielsen, 2012).

From the usability model accepted for this study, efficiency was measured by analysing the time it took to successfully complete a task (which relates to the metric "Time to activate one task" on the model), the number of errors as well as the number of failed commands issued by the participant (as seen in Figure 3.1). Specific metrics will be discussed in detail in Chapter 4.

Learnability

One of the most important aspects of a software program or interface in HCI is how intuitive it is to use. Learnability is defined by ISO 9126 as the effort that is required to learn how to utilise the application (ISO 9126, 2001). In essence, learnability is a measurement of how quickly a user can become proficient in using a product (Nielsen, 2012). During this study

learnability was investigated by requiring participants to complete multiple sessions. Differences in afore-mentioned metrics over time was utilised as an indicator of learning, which relates to the "Time to learn" metric as seen in Figure 3.1.

Satisfaction

Satisfaction can be seen as everything the user experienced during the interaction with the system (Nielsen, 2012). User satisfaction can be measured by using a post-test questionnaire for quantitative analysis (Lewis, 1995), or by interviewing the user for a more qualitative analysis (Akilli, 2005).

During this study, post-test questionnaires was used to gather data on user satisfaction in terms of the interfaces investigated. At the conclusion of the performance segment of each user test, the participants were asked to complete a written questionnaire asking them to rate their experience in using the specific device. These post-test questionnaires, which associates with the "Rating scale for users' satisfaction" metric (as seen in Figure 3.1), will be discussed in detail in Chapter 4.

Surveys were used together with user testing to gather primary data and will be discussed in the following section.

3.6. Surveys

As previously mentioned, this study made use of primary data in order to answer a causal question. Primary data can be collected by making use of surveys and have also been identified as being applicable to research in the field of HCI. Surveys are useful for answering research questions, finding solutions to problems that have been encountered as well as assessing needs and setting goals (Isaac & Michael, 1997). During this study, surveys were used to gather demographic data from the participants with the pre-test questionnaire as well as data on user satisfaction for each interface combination with the post-test questionnaire.

When making use of surveys, the response can either be structured or unstructured. While employing structured responses, collecting and summarising the results can be done more efficiently, however, they do restrict the responses from participants. These restrictions are due to the fact that participants are forced to choose the response from a list of pre-set options. Conversely, unstructured responses refer to responses that are in text format, where the

participant is allowed to answer the question in written form. Surveys can be broadly categorised into two types, namely questionnaires and interviews (Trochim, 2005).

3.6.1. Questionnaires

A questionnaire is a set of pre-planned written or printed questions, which is designed with the goal of gathering specific information from the research participant to answer a particular research question (Key, 1997).

Making use of a questionnaire has certain advantages as well as disadvantages. When using questionnaires, as opposed to interviews, the expense and time involved in training interviewers and sending them to interview participants are reduced. Since participants receive the same questionnaire, the data gathered from questionnaires will be better suited to comparing responses as all participants answered the same questions. However, the participants' motivation for their response is difficult to evaluate, thus affecting the validity of responses (Key, 1997). Two types of questions are available when using questionnaires (Key, 1997):

- **Open or unrestricted question**s – participants can respond in their own words, allowing for greater depth of response. Responses to open questions are difficult to interpret and summarise due to the open ended nature of the questions.
- **Closed or restricted questions** – participants can give a yes or no answer to the questions, give a short response, or check items from a list. This leads to responses that are relatively easy to interpret and summarise.

This study made use of closed questionnaires as one of the primary data collection methods which reduced costs as no additional interviewers had to be trained. It was also easier to compare and summarise the participants' responses. A pre-test questionnaire as well as a post-test questionnaire were administered. The pre-test questionnaire was used to gather demographic data as well as information on the participants' level of experience with computer games. The post-test questionnaire was used to gather data on the participants' experience of each of the user interfaces being investigated. A rating scale is usually employed during the design of closed or restricted questions. Rating scales force the participant to choose from a pre-selected set of answers. Several rating scales are available for use, such as the Likert, LPC and Guttman scales.

A Likert scale requires the participants to specify their responses on a rating scale where the two ends of the scale have opposite meanings (Olivier, 2004), for example, the response "Strongly agree" is on the one end and "Strongly disagree" on the other end of the scale (Tullis & Albert, 2013). An LPC scale (also known as a semantic differential (SD) scale) is comparable to a Likert scale with the exception that the respondents have to choose from a scale anchored on both sides by clear-cut responses (Olivier, 2004). The LPC scale measures participants' responses to words and concepts in terms of ratings on a bipolar scale, thus measuring the directionality of a response (good versus bad) and also the intensity of the response (somewhat to extremely) (Heise, 1970). When making use of a Guttman scale, participants are required to check each statement that they agree with. The statements are cumulatively structured, which means that if the participant chooses the last statement he would generally not only agree to the last statement, but to all the items above it as well (Trochim, 2005).

The Likert scale was used when designing the post-test questionnaires. This allowed the participants to indicate a positive or negative experience regarding the specific user interface as well as how usable they perceived the interface to be. The specific questionnaires that was used during this study will be discussed in Chapter 4.

3.6.2. Interviews

Interviews differ from traditional questionnaires as the interviewer completes the questions based on the participant's responses (Trochim, 2005). When making use of an interview, there will be a direct face-to-face attempt at obtaining reliable and valid data in the form of verbal responses from the participant. Put simply, it is a conversation between the researcher and the participant (Key, 1997).

For the purpose of this study the necessary data was collected by making use of user testing and closed question questionnaires, thus there was no need for interviews. In order to gather data that could be generalised to the wider population, a specific sampling method had to be used. Various sampling strategies and techniques will be discussed in the following section.

3.7. Sampling

A very important step in the research process is choosing a study sample because it is hardly ever practical, efficient or ethical to involve an entire population in the study. The population is the group of interest in terms of the study, thus the group to which the results of the study

should be generalisable to (Levy & Lemeshow, 2013). Although qualitative data was collected through the use of questionnaires, the majority of the data for this study was quantitative data gathered by means of user testing. The following discussion will focus on quantitative sampling approaches.

The goal of a quantitative sampling approach is to include a representative sample from the population, allowing for the results of the study to be generalised back to the population. The two main sampling methods used in research are probability sampling and non-probability sampling. The aim of this study will eventually influence and form the basis of the selection of an appropriate sampling method (Marchall, 1996). In the following two sections non-probability and probability sampling will be discussed in order to make a decision on the sampling methods for this study.

3.6.1. Nonprobability Sampling

Non-probability sampling does not involve random selection. There are several non-probability sampling strategies, namely quota, judgment and convenience sampling.

When using quota sampling, the researcher selects a sample where certain characteristics are present in proportion to their distribution in the overall population. Judgement sampling involves actively selecting the most useful participants, according to the researcher, in order to gather data that will answer the research question (Tansey, 2007). Due to time and monetary constraints, these two approaches did not lend themselves to the study at hand. The convenience sampling approach involves selecting the most accessible participants. It results in this approach being the most economical in terms of money, time and effort required (Marchall, 1996).

Convenience sampling was applied during the current study as the participants were students at the university where the study was conducted. The readily accessible participants were computer literate and the data gathered from these participants thus provided insight and possible answers to the identified research problem.

3.6.2. Probability Sampling

With probability sampling every individual in the population has an equal chance of being selected. This leads to results that are more likely to accurately reflect the entire population (Stehman, 1999).

There are four basic probability sampling techniques, namely cluster sampling, stratified sampling, systematic sampling and simple random sampling. When making use of cluster sampling, the entire population is separated into groups, referred to as clusters, and a random sample of these clusters are then selected (Trochim, 2005). Since this study did not make use of groups this sampling method is inappropriate and was not used. Stratified sampling techniques are generally used when the population is diverse, or where sub-populations can be set apart (Stehman, 1999). This sampling method was not utilised during this study as no unique groups were identified in the current study. When making use of the systematic sampling method individuals from the population are selected by utilising a random starting point as well as a fixed, periodic interval, referred to as the sampling interval (Trochim, 2005). This sampling method was not suitable to the current study as no sampling intervals were used and this sampling method was thus not utilised. The most straightforward form of random sampling is referred to as simple random sampling. Simple random sampling is a sampling method where the researcher selects a group of participants from the population. Each participant is selected totally by chance and each member of the population has an equal probability of being included in the sample. A simple random sample is intended to gather an impartial representation of a group to participate in a specific study (Stehman, 1999).

For the purpose of the study at hand the simple random sampling method was utilised as participants were drawn randomly from the convenience sample.

3.8. Conclusion

This chapter discussed the research design that was used in this study. The data required for the current study was gathered by utilising an experimental research design. Participants were assigned to the study using the simple random sampling method. User testing was conducted to gather data on the usability of the selected user interface combinations. Structured questionnaires making use of the Likert scale were also used to gather data on the participants' demographics as well as their satisfaction levels relating to the user interface combinations.

The following chapter will discuss the experimental design that was used to collect and analyse data in order to answer the research questions posed during this study.

CHAPTER 4 - Methodology

4.1. Introduction

The research design was discussed in the previous chapter where it was concluded that a comparative experiment had to be conducted, more specifically a usability evaluation involving user testing. Additionally, questionnaires were used to gather data pre- and post-test. During this chapter specific information relating to the methodology that was followed based on the selected research design will be discussed. This discussion will include the tools that were used as well as the tasks that had to be completed with the proposed interfaces.

4.2. Experimental Setup

A shooting genre game (Figure 4.1), developed using Microsoft XNA framework 4.0, was used for this experiment. While playing the game, the participant was required to position the cursor (which is shown as a crosshair) over the target (which is a flying duck) and then activate a command in order to shoot the target. Once a target had been eliminated, the participant had to activate another command in order to reload his/her weapon so that the following shot could be fired. If the participant missed the target, no reload was required as a reload was only required once a target had been eliminated.

Figure 4.1 - Duck Hunter

Therefore, the flow of game events proceeded as follows:

1. The target appeared where-after the participant had to reload and then fire any number of shots.

2. Once the target had been struck it was eliminated and disappeared from the game window.
3. The next target appeared after the previous target had been eliminated, but before the new target could be shot the reload action had to be triggered again.

Making use of this game was sufficient for the scope of this study as all the necessary data relating to the different user interfaces could be gathered for comparison. Three tasks were decided upon and included in the development of the game. Each task had a unique pattern for target starting positions. This pattern was repeated until all of the targets had been presented to the participant. The need to test stationary and moving targets was identified since these are typical gaming requirements. Therefore, the first task made use of stationary targets and the second and third of moving targets.

The first task required the participant to shoot at stationary targets that appeared, one at a time, in all four corners of the game window. The current target first had to be eliminated before the next target appeared. The first target appeared in the top-left corner of the game window, followed by the second target in the top-right, the third in the bottom-left and finally, the fourth target in the bottom-right corner. This pattern was repeated until the participant had successfully eliminated 20 targets.

The targets in task 2 had set starting locations, either on the left or right of the screen, and moved directly across the screen to exit on the opposite side. Screen sides alternated between targets, starting with the first target appearing on the right. The next target only appeared once the current target had been eliminated or if the current target exited the opposite side of the screen before it was eliminated. Consequently, after repeated exposure, participants were able to predict the starting location and direction of movement for this task.

In order to assess the interfaces with random unpredictable movement, which is indicative of a gaming environment, the last task implemented random starting locations and direction of movement. The starting positions and movement were pre-set, but appeared entirely random to the participant. This was to ensure that all participants were exposed to the same level of difficulty during this task. Thus, both predictive as well as non-predictive targets were included during testing. This also helped to clarify whether improvements in performance could be attributed to participants predicting the next target or whether the improvement was due to the participant being able to make better use of the interface as time went by.

The participants were expected to complete all three tasks, one with stationary targets and two with moving targets. Three interface combinations were utilised by each participant to complete the discussed tasks. These interfaces will be discussed in detail in the following section.

4.3. Interface combinations

For the purposes of this study two NUIs were compared to the traditional keyboard and mouse combination. The NUI technologies that were investigated included the Emotiv BCI and the Peregrine gaming glove. These technologies were combined in order to create two different NUI combinations that allowed for pointing as well as the activation of commands. For this game only two keys were used for two different actions. The first action was the reload action and could be activated by pressing the R key, which corresponds to the starting letter of the action. The second action was the shoot action that could be activated by pressing the S key, which also corresponds to the first letter of the action. Thus, R for reload and S for shoot. This was decided upon as many games make use of the first letter of an action to choose the key for that action. Therefore, both actions could be activated by using one hand, eliminating the need to move the hand responsible for mouse operation to the keyboard, which is a cause of long travel times between the mouse and keyboard.

Natural user interface 1

For the first NUI setup the Emotiv BCI, as seen in Figure 4.2, was used for testing, using the built in accelerometer as a method of pointing and then assigning certain keys to facial expressions for command activation.

Figure 4.2 - Emotive Brain Computer Interface (Emotiv Inc, 2015)

The cursor was controlled and moved by using small head movements, for example, tilting the head forward moved the cursor down, and tilting the head backwards moved the cursor up. Moving the cursor left and right was done by turning the head left or right. These movements are synonymous with moving one’s head to view a 3D scene therefore it was thought to be natural for the participants to move their heads in this manner.

The sensitivity level of the head movement could be adjusted through the use of the BCI software as seen in Figure 4.3. The interface therefore had the potential to be customisable, including when used within a gaming environment. During the pilot study different sensitivity levels were investigated and the most appropriate level was selected for usability testing.

Figure 4.3 - Emotiv mouse emulator

Using the BCI software a selection of keys and mouse clicks can be assigned to facial expressions (Figure 4.4), for example, the left click can be assigned to blinking the left eye, and the right click can be assigned to blinking the right eye. Other facial expressions such as the clenching of one's teeth, lifting the eyebrows, smirking left or right or even smiling can also be assigned to different keys.

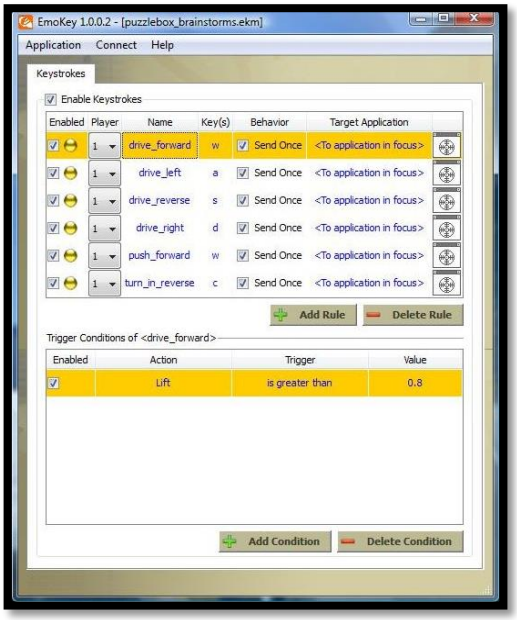

Figure 4.4 – EmoKey application

The disadvantage of this user interface is that all keys cannot be covered using facial expressions. This, however, was not a hindrance in this study as the game used for this study did not require the use of all the keys.

During this study, when using the BCI, the shoot command was activated by winking and the reload command was activated by raising the eyebrows. These two facial expressions were chosen as they are known to most people, therefore being more intuitive. Other facial expressions, such as clench and smirk, are not that well known and may have posed problems as participants may not have had a clear idea of what was required from them.

Natural user interface 2

The second NUI setup combined the accelerometer of the BCI for pointing with the Peregrine gaming glove (as seen in Figure 4.5). The cursor was controlled and moved with small head movements using the BCI accelerometer, while the actions were triggered by making use of the Peregrine glove.

Figure 4.5 – The Peregrine gaming glove (Peregrine, 2013)

Using the Peregrine glove software called GloveBox, all keys can be assigned to a certain part of the glove. For example, the letter A can be assigned to the tip of the index finger (Figure 4.6). When the tip of the thumb is pressed against the tip of the index finger the assigned key is activated. The tip of the index finger will be referred to as the touch point and the tip of the thumb as the activator point. The Peregrine glove has 17 touch points and 3 activator points allowing for over 30 user programmable actions, as seen in Figure 4.6. Thus, the glove is customisable and the user can choose which keys to map to which location on the glove.

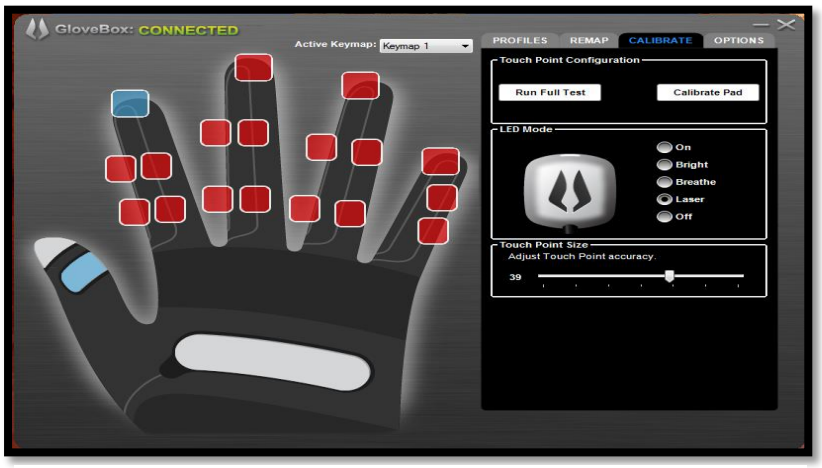

Figure 4.6 - Touch and activator points on Peregrine glove (Peregrine, 2013)

The advantage of this user interface is that all the keys can be assigned to a combination of touch points and activator points. A disadvantage is that users will have to remember which keys are assigned and how to access them. This was, however, not applicable to this study as

most 2D games only make use of a limited number of keys which will be easier for participants to remember.

During this study, the shoot command was allocated to the tip of the index finger as this simulates the pulling of a firearm trigger, and the reload action was activated by closing the hand, as this mimics the action of reloading a shotgun. The mimicking of certain actions have been proven to be a more natural method of input during gameplay (Logsdon, 2011). This made the actions more natural to the participant, consequently the actions might be more intuitive.

Each participant used both NUIs together with the traditional mouse and keyboard setup to play the game developed for this study. Making use of Nielsen's usability model (1994), data from all the user interfaces involved in the study was collected during user testing.

4.4. Usability metrics

Data collected from the three tasks allowed measurement of effectiveness, efficiency, satisfaction as well as learnability, as indicated in Nielsen's (1994) usability model.

Effectiveness was measured by the number of targets successfully eliminated during a task. The following data was utilised:

- **Target missed**: The total number of targets not eliminated during a task was recorded. Each target that left the game window before being eliminated contributed to this number. After each task the participant had a score out of 20 which was converted to a percentage. This metric could only be measured for tasks 2 and 3, where moving targets could move off-screen and result in a missed target.

Efficiency was measured by collecting data on the time spent on each active task and the number of errors. Thus, efficiency was measured using the following metrics:

Time Spent on task

- **Total time**: The total time the participant took to complete the specific task was recorded. Total time was analogous to "Time to achieve one task" metric as indicated on the usability model (Chapter 3.5.1) that guided the selection of the metrics for this study.

- Total time can be further deconstructed into the following measures:
    a. **Time for Reload**: The time it took the participant to activate the reload action after a target had been eliminated. This is a time measurement that influenced the total time of the task.
    b. **Initial contact time**: The time in milliseconds from when the target appeared till the participant positioned the cursor over the target.
    c. **Time on target**: The time that the cursor remained on the target before the shoot command was activated. This data excluded the time that the cursor spent off the target when an overshoot occurred. This metric did not accumulate and was consequently only relevant for the final time on the target – before the target was eliminated.

Errors

- **Miss**: When the participant triggered the shoot action while the cursor was not positioned over the target, the action was classified as a miss. The number of misses per target as well as per task were recorded.
- **Overshoot**: When a participant positioned the cursor over the target but then moved it off the target before shooting it, or if the target moved out from under the crosshair before being shot, the action was classified as an overshoot. The number of overshoots per target as well as per task were recorded.

Since learnability is defined as the effort that is required to learn how to utilise the application (ISO 9126, 2001), this metric was measured by requiring participants to complete a number of sessions over a period of time. In order to compare the different metrics over time the same participants were used in each session. During each session, participants had to complete all the tasks as defined for this study. Improvement of performance in terms of the metrics listed above indicated learnability. Participants were thus tested over 5 sessions in order to examine learnability.

In terms of questionnaires that measure satisfaction, there are a number of acceptable options, for example, the Usefulness, Satisfaction and Ease of Use Questionnaire and the System Usability Scale (SUS) (Sauro & Kindlund, 2005). For this study, the section of the Questionnaire for User Interaction Satisfaction (QUIS) that measures the overall satisfaction of the user was used (as seen in Appendix A). The QUIS was designed to measure participants'

personal satisfaction with specific aspects of an interface or device (Harper & Norman, 1993) that is relevant to this study and the reasoning behind the use of this particular questionnaire. Since the interfaces were essentially used for point-and-click tasks it was also deemed necessary to measure the users' satisfaction with specifically the pointing device as well. Therefore, the pointing device assessment questionnaire as described in ISO 9241-9 (ISO 9241, 2000) was used during the study. The Pointing Device Assessment Questionnaire (Appendix B) uses 13 questions and a 5-point Likert response scale to assess issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability (ISO 9241, 2000).

4.5. Pilot study

A pilot study was conducted in order to establish whether all the materials, protocols and documentation were adequately prepared. The pilot study preceded the test sessions and the participants for the pilot study were selected from a group of lecturers and post-graduate students. The pilot study focussed on the various sensitivity settings of the head-mounted accelerometer in order to select the optimal setting to be used in the actual study. The Emotiv BCI mouse emulator has customisable sensitivity settings ranging from low to high sensitivity. A high sensitivity requires smaller movement of the head in order to move the cursor a set distance than with low sensitivity. During the pilot study it was determined which of the sensitivity settings was the most effective to use during user testing. The three settings that were investigated were low, medium and high sensitivity (the actual settings were the minimum, middle and maximum of the sensitivity range). The pilot study was also conducted in order to test the methodology for the actual user testing. Therefore, the same methodology as discussed for the actual study was used. The pilot study will be discussed in detail in Chapter 5.

4.6. User Testing

4.6.1. Population, sampling and sample size

Computer literacy and a working knowledge of computers was a prerequisite for participants taking part in the study. Participants were selected from undergraduate students majoring in Computer Science at the University of the Free State in Bloemfontein. An open invitation to participate were addressed to all students. No specific criteria were used to filter participants as students majoring in Computer Science was considered computer literate. Therefore, the

sample was a random, convenience sample as participants were selected from students who accepted the invitation to participate in the study.

According to Nielsen and Landauer (1993), and as briefly mentioned in Chapter 1, five participants should be enough to gain insights when doing usability testing. During their study they started with the truth that zero participants will give zero insights. They found that when collecting data from a single participant, you will learn almost a third of all there is to know about the usability of the specific design. When adding a second participant it will add new insight, but not as much as the first participant did. As you add more participants, you will gain fewer new insights because you will keep seeing the same results repeating (Nielsen & Landauer, 1993).

Not only do the participants need to be randomly selected, but also the tasks and interfaces assigned to each participant. In an experiment, the order in which conditions (condition referring to the interface and task combination) are presented to participants can actually affect their behaviour due to fatigue or other external factors (Sharma, 1975). This can be overcome by randomly presenting the conditions to the participants to prevent them from adversely influencing the results of the study.

A Latin square can be used to ensure that each condition is preceded and followed by every other condition an equal number of times. For the Latin square to be correctly used, the number of participants used in the study should be a multiple of the number of conditions (MacKenzie, 2012). The current study investigated three interfaces, while making use of three tasks, thus nine conditions was used. Clearly, because the number of conditions were uneven, it was impossible to meet the requirement that each condition must be preceded and be followed by every other condition an equal number of times. In order to solve this problem two Latin squares were used. The second Latin square was formed by reversing the order of each row of the first Latin square (Sharma, 1975). Therefore, no two participants did the tasks in the same order. This also changed the order in which the participants made use of the interfaces, which balanced the effect of learning, fatigue and interactions between reactions to various experimental conditions (Bailey, 2008; Hinkelmann & Kempthorne, 2007).

This resulted in 18 possible combinations of conditions (as seen in Appendix C), therefore, by selecting eighteen participants every combination was investigated, resulting in enough insight

gained into the technology to compare the different user interfaces and render an accurate verdict.

It is important to note that the sample size was small for several other reasons as well. First, a small sample size allows for longer, more in-depth usability testing (Nielsen, 2000). More time with the participants was necessary because getting used to the new user interfaces took some time. Training sessions were used to introduce participants to the different user interfaces that were investigated. Secondly, testing a small number of participants was more reasonable in terms of logistics and scheduling. Test sessions took place on a one-to-one basis, a session with each participant lasted up to 45 minutes per session.

4.6.2. Test sessions

The test sessions took place in a laboratory environment where the facilitator was present at all times, without interfering with the participant's tasks. The instructions that were given to the participants were predefined and written in a document (Appendix D for protocol).

During the first session, each participant was required to complete a pre-test questionnaire (Appendix E), which gave an indication of the experience level that the participants had with gaming. Additionally, each participant's current level of comfort using personal computers and the technology used in the study was extrapolated from the collected data. It was expected that users who were more familiar with the selected NUI, would be responsible for significantly fewer errors and able to learn how to effectively use the interfaces more quickly than those who were not as familiar with the selected NUI. The pre-test questionnaire also included questions related to demographic information.

Every participant was asked to participate in each of the 5 test sessions. There was one session per week for each participant, with a specific time scheduled on the same day of the week, resulting in each participants' session being spaced similarly to every other participant. Three tasks per interface were performed during a test session, which took no longer than 45 minutes. In order to avoid order effects such as learning, the user interfaces were not used in a set order but were randomised using the balanced Latin square (Appendix C), as discussed in Section 4.6.1.

Each task, as discussed in Section 4.2, was performed using the traditional keyboard and mouse, the BCI combination as well as the BCI and Peregrine glove combination. As previously mentioned the tasks were the following:

- Task 1: Twenty stationary targets were presented one after the other to the participants. The participants then had to attempt to eliminate all of the targets by making use of the selected interface.
- Task 2: Twenty predictable moving targets were presented one after the other to the participants. These targets moved from left to right and from right to left across the game window. The participants then had to attempt to eliminate all the targets by making use of the selected interface.
- Task 3: Twenty targets were presented to the participants. In this case the targets were moving and appeared and moved in an unpredictable manner.

The tasks were explained and demonstrated to the participants. They were asked to work as fast as possible while still aiming for high accuracy. After completing a user interface task, the participants rested for two minutes before receiving instructions on the next task.

4.7. Issues of reliability and validity

Validity was established by using existing questionnaires to gather data. The QUIS questionnaires (Harper & Norman, 1993), in conjunction with the Pointing Device Assessment Questionnaire (ISO 9241, 2000), was used during this study.

4.8. Data analysis and interpretation

Descriptive and inferential statistics were used to analyse the data that was gathered. Descriptive statistics included mean and standard deviation which gave an indication of the spread of the data set and the relationship of the mean to the rest of the data. Inferential statistics was used to test the stated hypotheses and infer the results back to the general population. These tests were determined by the collected data. The same participants were included in every session, therefore the repeated measures ANOVA was used to analyse the data (Field & Hole, 2002). Where significant differences between variables were detected with the repeated measures ANOVA, a Wilcoxon post-hoc test using the Bonferroni correction was conducted to determine which interfaces or sessions were responsible for the significant difference. If the

ANOVA indicated that there was significant interaction between the two factors, namely session and interfaces, separate ANOVAs and Friedman tests were conducted.

The specific statistical tests used for analysis will be discussed in detail in Chapter 6.

4.9. Confidentiality and ethical issues

Participants were assured of anonymity and was free to leave at any time should they wish to do so. All data and data analysis will remain anonymous. Each participant signed an informed consent form, which formed part of the pre-test questionnaire as seen in Appendix E.

Since people were involved in the testing, ethical clearance was obtained from the ethics committee of the University of the Free State prior to conducting the study (Appendix F).

4.10. Conclusion

This chapter discussed the methodology that was followed during the investigation. This discussion included the experimental setup that was used, the interface combinations that were utilised during user testing as well as the sampling used for the study.

The pilot study was conducted to reveal the optimal sensitivity setting for the head-mounted accelerometer that formed part of each NUI in the actual study. In the next chapter the analysis of the data gathered from the pilot study will be discussed.

Chapter 5 – Pilot study

5.1. Introduction

The previous chapter discussed the methodology that was followed during this study. The discussion included the tool that was used, the tasks the participants were required to complete and the metrics that were analysed in order to establish the usability of each interface.

A pilot study was conducted to establish whether all tools, protocols and documentation were adequately prepared. In addition, the pilot study was a means of determining which setting was ideal for use within the context of the study. The pilot study preceded the test sessions and was conducted according to the actual test protocol and followed the same methodology as was discussed in the previous chapter. This chapter will briefly discuss the findings of the pilot study.

5.2. Pilot study analysis

The Emotiv BCI, which was used for cursor control, is equipped with a sensitivity scale. The level of sensitivity is linked to the distance that the head has to turn in order to move the cursor. If the sensitivity is set on high, small movements are required to move the cursor a certain distance, while the head has to be turned further to move the cursor the same distance when using the low sensitivity setting. Since the setting can have an effect on the usability of the BCI, it is important that an optimal setting is chosen for the testing. A pilot study provides a means to determine which setting is optimal for the majority of participants. Therefore, a pilot study was conducted prior to the actual user testing to investigate which sensitivity setting would assure optimal cursor control, but also to test the protocol and questionnaires of the main study. The pilot study was designed to determine which of these settings was optimal for cursor control in a gaming environment. Twenty participants were required to complete tasks 1 and 2 on each sensitivity level. Participants for the pilot study were selected from a group of lecturers and post-graduate students who were well informed about usability testing protocols and methods which that allow them to give valuable feedback in terms of the protocol and methodology while still allowing the optimal setting to be determined.

During the pilot study all the metrics identified as relevant to test the usability of the interfaces were captured to ensure the correct capturing of the data for the main test. However, since the focus of the pilot study was on the sensitivity setting for the head-mounted accelerometer, only

certain metrics were analysed, namely the total time, first contact time, time on target and the number of overshoots. These metrics were chosen since they provide a strong indication of cursor control.

Analysis was conducted by using the Friedman tests to investigate whether there were any significant differences present. In the case of significant differences Wilcoxon post-hoc tests were applied in order to find the specific differences. Significance levels of 0.05 were used throughout, but where relevant highly significant results were obtained, these will be reported as such. In the case of the Wilcoxon post-hoc tests, the Bonferroni adjustment was applied by taking the significance level that was initially used (0.05) and dividing it by the number of post-hoc tests. Therefore, the significance level of 0.05/3 = 0.017 was used when applying the Bonferroni adjustment to the Wilcoxon post-hoc.

5.2.1. Total time

The total time to complete the task was measured from when the task was started to when the task was completed by the participant.

The following hypothesis was formulated for the analysis of the data:

- $H_{0,1}$: The sensitivity setting used has no effect on the time taken to complete a task.

Friedman tests showed a statistically significant difference between the different settings for both tasks at an α-level of 0.01 – Task 1: $\chi^2(2) = 10.000$, $p < 0.01$, Task 2: $\chi^2(2) = 17.200$, $p < 0.01$.

Therefore, the null hypothesis $H_{0,1}$ can be rejected for task 1 and 2. This result indicates that there was a significant difference in the time it took to complete both tasks when making use of the different sensitivity settings. The table below summarises the Wilcoxon post-hoc signed rank test between each pair of interfaces:

Table 5.2.1.1 - Wilcoxon post-hoc: Total time

	Low - Med	**Low – High**	**Med - High**
Task 1	Z = -2.165, p > 0.017	Z = -3.099, p < 0.017*	Z = -1.605, p > 0.017
Task 2	Z = -1.419, p > 0.017	Z = -3.454, p < 0.017*	Z = -2.651, p < 0.017*

*** = p < 0.017**

As can be seen clearly from Table 5.2.1.1, there was a significant difference between the low and high sensitivity level for task 1, as well as a significant difference between the low and high, and the medium and high setting during task 2. Inspection of the mean times (Table 5.2.1.2) to complete task 1 indicated that the low sensitivity had faster average completion times than the medium, but significantly faster times than the high sensitivity setting. Inspection of the mean times for task 2 reveals that the low sensitivity setting was significantly faster when compared to both medium and high sensitivity.

Table 5.2.1.2 - Mean times for the total time metric: Task 1 & 2

	Low	**Medium**	**High**
Task 1	54.5	59.6	65.7
Task 2	50.1	55.3	66.0

Therefore, it could be concluded that making use of the low sensitivity setting to complete tasks, where stationary as well as moving targets will be encountered, will result in faster completion times.

5.2.2. Initial contact time

The initial contact time was measured from when the target was presented until the cursor first made contact with the target.

The following hypothesis was formulated for the analysis of the data:

- $H_{0,1}$: The sensitivity setting used has no effect on the time taken to make the initial contact with the target.

Friedman tests showed a statistically significant difference between the different settings for one of the two tasks at an α-level of 0.05 – Task 1: $\chi^2(2) = 1.200$, $p > 0.05$, Task 2: $\chi^2(2) = 9.100$, $p < 0.05$. Thus, the null hypothesis $H_{0,1}$ can be rejected for task 2 but cannot be rejected for task 1. This result indicates that there is a significant difference in the first contact time when using the three settings for task 2. A summary of the Wilcoxon post-hoc signed rank test between each pair of interfaces is presented in Table 5.2.2.1:

Table 5.2.2.1 - Wilcoxon post-hoc: First contact time

	Low – Med	**Low – High**	**Med – High**
Task 2	Z = -2.016, p > 0.017	Z = -0.709, p > 0.017	Z = -2.203, p > 0.017

*** = p < 0.017**

Using the conservative Bonferroni correction, the alpha-level is 0.017. However, at that level no pairwise comparisons were found to be significantly different. Nevertheless, at a less conservative alpha value, the low–medium and medium–high sensitivity levels were significantly different ($p < 0.05$). Inspection of the mean first contact times (Table 5.2.2.2) during task 2 indicated that the medium sensitivity had significantly faster times, at the less conservative alpha-level, than the low as well as the high sensitivity.

Table 5.2.2.2 - Mean times for the first contact time metric: Task 2

	Low	**Medium**	**High**
Task 2	37.7	35.5	39.5

Therefore, it could be concluded that making use of the medium sensitivity to make first contact with the target would be the better option, although the low sensitivity would also be a viable choice if the opportunity presented itself.

5.2.3. Time on target

The time on target metric was measured from when the cursor was positioned over the target until the target was eliminated on condition that the cursor did not leave the target during this time.

The following hypothesis was formulated for the analysis of the data:

- $H_{0,1}$: The sensitivity setting used has no effect on the time spent on the target before it is eliminated.

Friedman tests showed a statistically significant difference between the different settings for one task – Task 1: $\chi^2(2) = 0.300$, $p > 0.05$, Task 2: $\chi^2(2) = 7.900$, $p < 0.05$. Consequently, the null hypothesis $H_{0,1}$ can be rejected for task 2 but not task 1. This result indicates that there is a significant difference in the time spent on the target when using the different sensitivity levels with moving targets. The table below summarises the Wilcoxon post-hoc signed rank test between each pair of interfaces:

Table 5.2.3.1 - Wilcoxon post-hoc: Time on target

	Low – Med	**Low – High**	**Med - High**
Task 2	$Z = -0.336$, $p > 0.017$	$Z = -2.091$, $p > 0.017$	$Z = -2.632$, $p < 0.017$*

*** = p < 0.017**

As can clearly be seen from Table 5.2.3.1, there was a significant difference between the medium and high sensitivity setting in terms of the mean time on target during task 2. Inspection of the mean times on target (Table 5.2.3.2) during task 2 indicated that the high sensitivity setting had faster times than the low as well as the medium setting.

Table 5.2.3.2 - Mean times for the time on target metric: Task 2

	Low	**Medium**	**High**
Task 2	4.2	4.2	3.7

Therefore, it could be concluded that making use of the high sensitivity setting when targeting and eliminating a target will result in the least time spent on the target before it is eliminated. However, when using the high sensitivity setting it was observed that the participants were struggling to make small adjustments with their heads in order to keep the cursor positioned over the target. It was more problematic for moving targets as they would have to realign the cursor in the path of the oncoming target in order to ambush it. The smaller corrections were easier with stationary targets.

It may be reasoned that this may not be a positive result for the high sensitivity setting as it indicates that the participants were not able to actively track the target but would rather try to ambush the moving target. This ambush tactic would result in less time spent on the target before it was eliminated but would only work in settings where the path of the target was predictable, which is not the case in most games. This was not the case with the low and medium setting where the head movement required in order to actively track the target was not as small as with the high sensitivity setting. Therefore, in light of the difficulty experienced with correcting and positioning, even though the high setting has the lowest time, it is recommended that either the medium or the low setting is a more viable option based on this metric.

5.2.4. Number of overshoots

The number of overshoots metric was measured by totalling the number of times the participant positioned the cursor over the target and then lost contact with the target without eliminating it. The following hypothesis was formulated for the analysis of the data:

- $H_{0,1}$: The sensitivity setting used has no effect on the number of overshoots.

Friedman tests showed a statistically significant difference between the different settings for both tasks – Task 1: $\chi 2(2) = 14.256$, $p < 0.01$, Task 2: $\chi 2(2) = 20.100$, $p < 0.01$. Thus the null hypothesis $H_{0,1}$ can be rejected for task 1 and 2. This result indicates that there is a significant difference in the number of overshoots when using the three sensitivity settings. The table below summarises the Wilcoxon post-hoc signed rank test between each pair of interfaces:

Table 5.2.4.1 - Wilcoxon post-hoc: Number of overshoots

	Low - Med	**Low – High**	**Med - High**
Task 1	Z = -3.382, p < 0.017*	Z = -3.288, p < 0.017*	Z = -1.794, p > 0.017
Task 2	Z = -2.782, p < 0.017*	Z = -3.585, p < 0.017*	Z = -1.345, p > 0.017

*** = p < 0.017**

As can clearly be seen from Table 5.2.4.1, there was a significant difference between the low and medium as well as the low and high setting during tasks 1 and 2. When inspecting the means for tasks 1 and 2 it becomes clear that the low setting was responsible for significantly less overshoots than the medium and high setting (Table 5.2.4.2).

Table 5.2.4.2 – Mean for the number of overshoots metric: Task 1 & 2

	Low	**Medium**	**High**
Task 1	15.4	23.7	28.0
Task 2	16.1	24.7	29.4

It can be argued that individuals rather make use of their eyes to scan the details in the environment in front of them. The head is only turned to move the eyes in the direction of a new location where the eyes again will perform the smaller focusing movements. This can explain the participants' difficulty with the high sensitivity setting as these small movements are usually performed by the eyes and not by the head. This resulted in participants rather moving the cursor in the path of the target and waiting for it to intersect with the cursor before activating the shoot command.

Therefore, it could be concluded that making use of the low sensitivity setting when targeting and eliminating both stationary and moving targets will result in the least number of overshoots. This means that it is easier to position and keep the cursor stable over the target when using the low sensitivity setting, which is a desirable trait in a primarily point-and-click GUI.

5.2.5. Anecdotal observations

Some additional observations and alterations were made during the pilot study which were incorporated into the methodology for the actual usability test. This includes difficulty experienced with the head controlled cursor as a result of the cursor being bound to the game window, which is the case in most games. Thus, when a participant turned his/her head too far to the left or right (or top and bottom), the cursor stopped at the edge of the game window. This implies that the cursor does not necessarily move the same distance as the head movement. The cursor is then off-centre when the participant turns his/her head back to face the middle of the game window, which makes accurate aiming very difficult. In order to correct for this phenomena an additional command was added to the game which could be activated at any stage when the participant felt that the cursor was no longer aligned correctly. The reset command could be activated when the participant was satisfied that his/her head was aimed at the middle of the game window and the cursor was not, with the result that the cursor would be repositioned to the centre of the game window. During the pilot study this solution worked very well, with participants quickly learning when to use the reset command, as well as how to perform the reset in a very short time. Thus, for the usability study the reset command was bound to the space bar key when making use of the keyboard, the tip of the middle finger when using the glove and clenching the teeth when utilising facial expressions.

Additionally, it was observed that the time spent to complete a session was too long, as it regularly went pass the hour mark. This was not practical due to the fact that 5 sessions per participant had to be completed for this study. Each session included 20 targets per task and the participants had to perform 3 tasks per interface combination. During the analysis of the data it was confirmed that the loss of data would be negligible if the target count was reduced to 15 in order to save some time per session. Thus, although the tasks will remain unchanged, it was decided that 15 instead of 20 targets were to be used during each task.

5.3. Conclusion

When inspecting the results from the individual metrics, it is clear that the low sensitivity setting was found to be the best in 2 of the 4 metrics, namely the total time and number of overshoots. The medium setting was found to be the best in terms of first contact time. However, this was only the case during task 2, with task 1 showing no significant differences between the sensitivity settings. The high sensitivity setting was found to be responsible for

the least time spent on the target before it was eliminated. This, however, was only in the case of task 2, with no significant differences found for task 1.

When inspecting the number of overshoots for the high sensitivity setting it becomes clear that participants experienced greater difficulty with cursor control than with the low and even medium sensitivity, which strengthens the argument that the high sensitivity is responsible for the loss of accurate control of the cursor. Taking into account that the high sensitivity is therefore not recommended for use, but that low and medium had the same average for this metric, one could suggest that the low setting is the best setting in 3 of the 4 metrics and was consequently utilised during this study. This low sensitivity makes better use of the natural head movement that individuals use on a daily basis, consequently larger movements of the head will be more natural than small adjustments, which will have to be learned.

The next chapter will discuss the results of the user testing.

Chapter 6 - Analysis

6.1. Introduction

The analysis of the data gathered from the pilot study was performed and discussed in the previous chapter. From the analysis, it was concluded that the low setting (the actual setting was the minimum of the sensitivity range) is the optimal setting for use in 2D gaming. The analysis of the full study will be discussed in this chapter and will include the analysis of objective as well as subjective data. The statistical analysis of the different metrics will include data for three tasks over 5 sessions. Data gathered during user testing will be used to discuss the effectiveness, efficiency and learnability of the three interfaces, whereas data from questionnaires will be used to discuss the users' level of satisfaction with the interfaces.

6.2. Analysis

The experiment was conducted over several sessions, with each session following the same protocol and including the same tests. The same participants were included in every session, therefore the repeated measures ANOVA will be used to analyse the data (Field & Hole, 2002).

6.2.1. Repeated measures ANOVA

The repeated measures ANOVA assumes that the dependent variable is measured at the continuous level and that the independent variable consists of at least two related groups. There should also be no significant outliers in the two or more related groups (Minke, 1997).

Another assumption is that the distribution of the dependent variable in the two or more related groups is normally distributed (Minke, 1997). For this study the time measurements were in milliseconds and there were a number of instances in which the data failed the test for normality. In order to counter this, the time measurement was converted to 1/time as done in previous studies (cf. Beelders, 2011; Dollman, 2014). Although it is an assumption, the repeated measures ANOVA is robust for violations of normality (Keselman, Algina & Kowalchuk, 2001). Thus, if there are violations of normality, it will not have an impact on the analysis of the data and the repeated measures ANOVA will still be used for the analysis. Regardless, the data was tested for normality in all instances, but for the sake of brevity normality results will not be reported here.

The last assumption is sphericity**,** which indicates that the variances of, and the correlations among the repeated measures are all equal (Vasey & Thayer, 1987). Mauchley's sphericity test will be used to verify whether the assumption of sphericity is met. When sphericity is violated, the results will be reported using the Greenhouse-Geisser adjustment (Abdi, 2010).

In the case of significant differences between variables, a Wilcoxon post-hoc test using the Bonferroni correction will be conducted to determine which interfaces or sessions were responsible for the significant difference. When the ANOVA indicates that there is significant interaction between the two factors, namely session and interfaces, separate analyses will be conducted by isolating each factor in turn, thus separate ANOVAs and Friedman tests will be conducted. Tasks will be grouped and discussed together when there is no significant interaction between the two factors, namely interface and session. In the case where there is significant interaction between the two factors, each task will be discussed separately as the discussion contains more detail in this instance.

6.2.2. Analysis of separate tasks

During all three tasks the participants had to eliminate 15 targets by positioning the cursor over the target and activating the shoot command. For each task, as described in detail below, a total of 15 targets appeared at different locations. Stationary targets remained onscreen until the participant effectively eliminated it by successfully positioning the cursor over the target and activating the shoot command. Moving targets had to be eliminated by the participant but could also move off-screen.

As discussed in Chapter 4:

1. Task 1 constituted stationary targets that appeared in a sequence, namely top-left, top-right, bottom-left, bottom-right corner of the game window. This pattern repeated until 15 targets had been presented.
2. For task 2, targets moved from one side of the game window to the other, alternating between left-to-right and right-to-left movement in the vertical middle of the game window.
3. During task 3 targets appeared at predefined locations along the game window and moved to the opposite side of the game window. The targets are referred to as predefined random targets as their starting locations were pre-selected, but would

appear random to the participants. Thus, all participants will have the same starting position for all targets.

6.2.3. Interaction techniques

For the purposes of the analysis, the interaction technique will be referred to as an interface. The interfaces used are:

1. keyboard and mouse (KM),
2. an Emotiv BCI head-mounted accelerometer in conjunction with the Peregrine gaming glove (BCIG),
3. an Emotiv BCI head-mounted accelerometer in conjunction with the Emotiv BCI facial expressions (BCIF).

During tasks 1, 2 and 3 data was collected for several metrics as discussed in Chapter 4. The analysis for each metric is discussed below.

6.3. Total time to complete the task

The total time to complete the task was measured from when the task was started to when the task was completed by the participant. The total time is comprised of the sum of the individual time intervals it took the participant to eliminate 15 targets.

The following hypotheses were formulated for the analysis of the three tasks:

- $H_{0,1}$: The interface used has no effect on the time taken to complete a task.
- $H_{0,2}$: There is no difference in the time taken to complete the task between multiple sessions.

6.3.1. Task 1 and 2: Total time to complete task

In this instance, the assumption of sphericity was met for both task 1 (χ^2 (9) = 6.130, $p > 0.05$) and task 2 (χ^2 (9) = 8.885, $p > 0.05$) at an α-level of 0.05.

Interfaces

Since there is no significant interaction between the two factors, tasks 1 and 2 will be discussed together. At an α-value of 0.05 the null hypothesis $H_{0,1}$ can be rejected for both task 1 (F (2, 51) = 153.557, $p < 0.05$) and task 2 (F (2, 51) = 119.619, $p < 0.05$), therefore the interface used has an effect on the time taken to complete the task for both stationary and moving targets. The

results of the post-hoc tests are shown in Table 6.3.1.1 with task 1 on the first line of each cell and task 2 on the second line. It is important to note that the cell where two interface intersect relates to the comparison of the two interfaces.

Table 6.3.1.1 - Summary of results for the post-hoc tests: Total time – Task 1 & 2

	KM	**BCIG**
BCIG	t(17) = -11.144* t(17) = -7.921*	
BCIF	t(17) = -17.287* t(17) = -15.468*	t(17) = -6.143* t(17) = -7.542*

*** p < 0.05**

As can clearly be seen from Table 6.3.1.1, there was a significant difference between all three interfaces in terms of the time taken to complete task 1 and task 2. Inspection of the mean times to complete task 1 indicated that KM (22.0 seconds) had significantly faster times than BCIG (35.7 seconds) and BCIF (54.9 seconds). The same trend was seen for task 2, where KM (22.5 seconds) had significantly faster times than BCIG (32.6 seconds) and BCIF (60.4 seconds).

Based on this, together with the fact that post-tests indicate a significant difference between all interfaces, it could be concluded that target acquisition and elimination of stationary or predictable moving targets when using KM is significantly faster than when using BCIG and BCIF. Therefore, KM, which is known to the participants, is better suited to acquiring and eliminating a stationary or predictable moving target within a 2D game and can be recommended for use in this regard.

Although KM is better suited to the specific task, a significant difference in the mean times between BCIG and BCIF was also noted. When having to choose between these two interfaces BCIG would be the more suitable interface to use.

Sessions

At an α-value of 0.05 the null hypothesis $H_{0,2}$ can be rejected for task 1 ($F(4, 204) = 53.576$, $p < 0.05$) and task 2 ($F(4, 204) = 58.152$, $p < 0.05$), therefore, there is a significant difference in the time spent to complete tasks 1 and 2 from one session to the next. The results of the post-hoc tests in the table below, with task 1 results on the first line of each cell and on the second line for task 2 (Table 6.3.1.2).

Table 6.3.1.2 - Summary of results for the post-hoc tests: Total time sessions

	Session 1	**Session 2**	**Session 3**	**Session 4**
Session 2	t(17) = 6.878* t(17) = 6.148*			
Session 3	t(17) = 8.737* t(17) = 9.525*	t(17) = 1.913 t(17) = 2.729		
Session 4	t(17) = 10.524* t(17) = 9.954*	t(17) = 4.533* t(17) = 3.843*	t(17) = 2.673 t(17) = 1.509	
Session 5	t(17) = 13.968* t(17) = 13.19*	t(17) = 7.844* t(17) = 9.875*	t(17) = 4.847* t(17) = 6.353*	t(17) = 2.679 t(17) = 4.17*

*** p < 0.05**

As can clearly be seen from Table 6.3.1.2, there was a significant difference between most sessions for tasks 1 and 2. There was a steady decline in the time it took to complete tasks 1 and 2 from sessions 1 to 5 (Chart 6.3.1.1 and Chart 6.3.1.2).

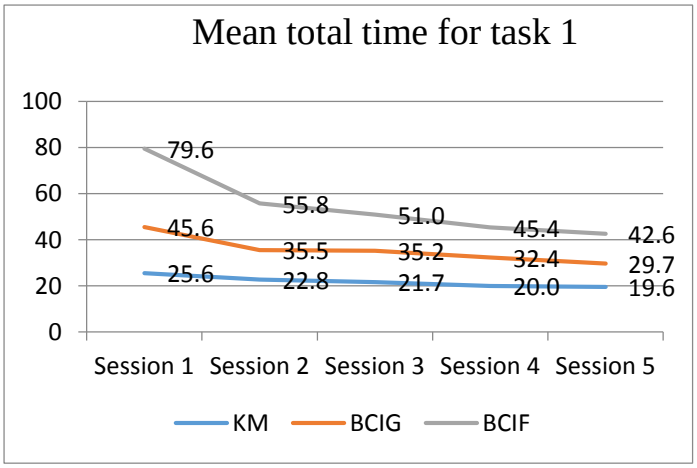

Chart 6.3.1.1 – Mean times for task 1

Inspection of the mean times (shown on the charts by each session) indicated that the completion time during session 5 was the fastest for both tasks. This indicates that learning took place for all interfaces. Part of the improvement was due to the participant becoming more familiar with the game, since improvement was present when using the keyboard as well – an

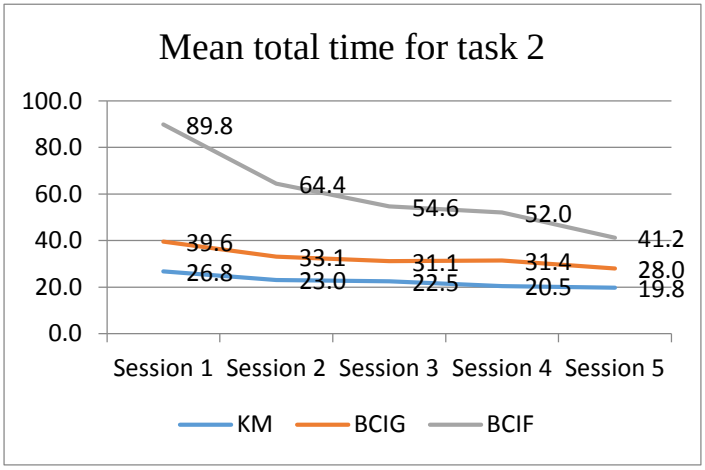

Chart 6.3.1.2 - Mean times for task 2

interface with which the participants were familiar with. When inspecting the means for sessions 1 and 5 for all interfaces and for both tasks it appears as though most improvement in time was achieved for BCIF.

KM, which participants use on a regular basis, had a 6 second improvement on the mean time from sessions 1 to 5 for task 1, and a 6.9 second improvement for task 2. Comparing this to BCIG, (improvement for task 1 - 15.9 seconds and task 2 - 11.6 seconds), as well as BCIF which showed large improvements (task 1 - 37.1 seconds and task 2 - 48.6 seconds), it becomes clear that these are improvements with the interfaces and not with the game, as KM did not show similar levels of improvement. It would be interesting to determine whether additional sessions would have seen the trend continuing and where the improvement would eventually level out. Therefore, it could be established that repeated use of the interfaces yielded improved completion times for all interfaces.

6.3.2. Task 3: Total time to complete task

The assumption of sphericity (χ^2 (9) = 56.951, $p < 0.05$) was not met at an α-level of 0.05. A within-subjects repeated measures ANOVA indicated that there was significant interaction between the two factors, namely session and interfaces ($F(4.989, 127.207) = 10.201$, $p < 0.05$), therefore separate analyses had to be conducted by isolating each factor in turn, thus separate ANOVAs and Friedman tests were conducted.

Interfaces

Friedman tests showed a statistically significant difference between the different interfaces for all sessions – Session 1: $\chi^2(2) = 28.444$, $p < 0.01$, Session 2: $\chi^2(2) = 33.371$, $p < 0.01$, Session 3: $\chi^2(2) = 36.000$, $p < 0.01$, Session 4: $\chi^2(2) = 36.000$, $p < 0.01$, Session 5: $\chi^2(2) = 32.444$, $p < 0.01$. Therefore, the interface used has an effect on the time taken to complete task 3, since the null hypothesis $H_{0,1}$ can be rejected for all three interfaces. Table 6.3.2.1 summarises the Wilcoxon signed rank test (with the Bonferroni correction), post-hoc tests for the Friedman test, for each session and between each pair of interfaces:

Table 6.3.2.1 - Summary of Wilcoxon post-hoc tests: Total time – Task 3

	KM - BCIG	KM - BCIF	BCIG - BCIF
Session 1	Z = -3.593, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.593, p < 0.017*
Session 2	Z = -3.464, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*

Session 3	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*
Session 4	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*
Session 5	Z = -3.549, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*

*** = p < 0.017**

Once again there was a significant difference between all three interfaces for all sessions in terms of the mean total time taken to complete task 3. Inspection of the mean times to complete task 3 over 5 sessions indicated that KM (24.7 seconds) had significantly faster times than BCIG (36.3 seconds) and BCIF (61.0 seconds). This reveals that BCIG and BCIF were far slower than KM, therefore KM is better suited to targeting and eliminating unpredictable moving targets. BCIG showed faster times than BCIF and will thus be a better option if a choice has to be made between the two interfaces for the task at hand.

Sessions

The data was found to be spherical for KM but not for BCIG and BCIF (Table 6.3.2.2). The results of the sphericity test and repeated measures ANOVA are tabulated below:

Table 6.3.2.2 - Summary of Mauchley's sphericity test and ANOVA

	KM	**BCIG**	**BCIF**
Sphericity test	χ^2 (9) = 9.554	χ^2 (9) = 31.971*	χ^2 (9) = 23.267*
ANOVA	F (4, 68) = 10.972*	F (1.912, 32.503) = 23.146*	F (2.452, 41.692) = 32.318*

*** = p < 0.05**

As can clearly be seen from the table, $H_{0,2}$ can be rejected for all interfaces, therefore the session has an effect on the total time to complete the task. Post-hoc tests were conducted to determine which sessions differed significantly from one another and are summarised below (KM is on the topmost row of each cell, BCIG directly below and BCIF on the last row of each cell):

Table 6.3.2.3 – Summary of post-hoc tests: Total time sessions

	Session 1	**Session 2**	**Session 3**	**Session 4**
Session 2	t(17) = -2.682 t(17) = -4.458* t(17) = -3.758*			
Session 3	t(17) = -4.831* t(17) = -5.38* t(17) = -5.701*	t(17) = -1.882 t(17) = -2.886 t(17) = -4.085*		
Session 4	t(17) = -4.197* t(17) = -5.279* t(17) = -5.807*	t(17) = -2.466 t(17) = -2.241 t(17) = -3.548*	t(17) = -1.151 t(17) = -0.283 t(17) = -0.536	
Session 5	t(17) = -5.215* t(17) = -6.303* t(17) = -8.372*	t(17) = -2.859 t(17) = -3.891* t(17) = -7.721*	t(17) = -2.018 t(17) = -1.325 t(17) = -4.365*	t(17) = -0.712 t(17) = -1.198 t(17) = -6.266*

*** p < 0.05**

It is clear from Table 6.3.2.3, that there are significant differences between session 1 and the other sessions, with the exception of KM for session 2. There was also a significant difference for all three interfaces between session 1 and session 5, but not between session 4 and 5 and sessions 3 and 5. Thus there was a large improvement from session 1 to 2, but smaller improvements thereafter. Inspection of the mean times (Chart 6.3.2.1) indicated that the total time during session 5 was the fastest. When inspecting the difference in the mean times for sessions 1 and 5 between BCIG and the traditional interface KM, the difference for session 1 was 19.5 seconds. This was reduced to only 8.8 seconds during session 5. A more drastic change was seen between the difference in the time to complete the task using KM and BCIF, which was reduced from 53.2 seconds during session 1 to 21.9 for session 5.

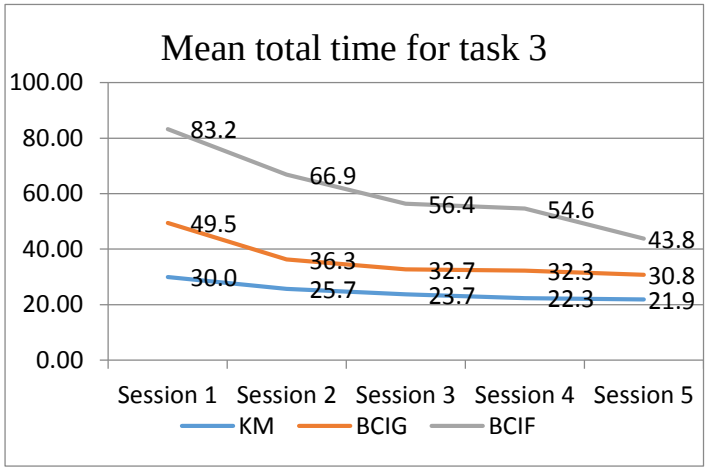

Chart 6.3.2.1 - Mean times to complete task 3

Thus, BCIG was initially 19.5 seconds slower than KM, but improved to the extent that it was finally only 8.8 seconds slower. BCIF improved from 53.2 seconds slower to 21.9 seconds slower over the course of the 5 sessions. Therefore, while KM improvement indicates increasing familiarity with the game, the greater improvement of BCIG and BCIF cannot solely be attributed to improvement with the game but is indicative of participants becoming adept in the interfaces.

Although the original use of BCIF was slow compared to KM, the rate of learning shows that BCIF could successfully be used for 2D gaming. However, the completion times for BCIG compared much better with KM, indicating that successful use of the interface could be achieved without much training, which further indicates that BCIG is more intuitive than BCIF.

6.3.3. Total time summary

During all three tasks KM proved to be the best interface to use when eliminating stationary, predictable as well as unpredictable moving targets. Since this is the interface with which the participants were the most familiar with, this result is understandable. BCIG presented the second best total time for all three tasks and BCIF proved to be the slowest of the three interfaces. It was also concluded that significant learning took place during all three tasks, which included learning in terms of the game itself but also and more importantly with the new interfaces.

Although BCIF was the slowest of the interfaces it was proven that it could be successfully used as an interface when playing 2D games. BCIF produced the best improvement times and this may be due to the participants never before having to make use of their head and face for computer input. At first, participants were unfamiliar with the workings of the interface and had to concentrate more when using it, thus it can be reasoned that this interface was not very intuitive to use. Even though it was not natural to use and was the slowest of the interfaces tested this interface opens up new possibilities for users who are unable to use their hands and arms.

Furthermore, daily use of a computer involves moving the cursor over items such as icons and then clicking a mouse button. This is similar to what participants were expected to do during task 1 that used stationary targets. Therefore, it can be reasoned that this interface may be of potential use for disabled individuals on a daily basis when performing routine tasks on a computer.

6.4. Reload time

The reload action had to be performed for each target before the shoot command could be activated. The reload time was measured from when the target was presented to the participant until the participant triggered the reload action. Participants could not reload before the target was visible.

The following hypotheses were formulated for the analysis of the three tasks:

- $H_{0,1}$: The interface used has no effect on the reload time.
- $H_{0,2}$: There is no difference in the reload time between multiple sessions.

6.4.1. Task 1, 2 and 3: Reload time

Since whether targets are stationary or moving has no influence on the reload of the weapon, the average of the three tasks (per interface and per participant) will be used for this analysis. The assumption of sphericity (χ^2 (9) = 0.026, $p < 0.05$) was not met at an α-level of 0.05. A within-subjects repeated measures ANOVA indicated that there was significant interaction between the two factors, namely session and interfaces ($F(2.914, 74.311) = 3.169$, $p < 0.05$), thus separate ANOVAS and Friedman tests were conducted.

Interfaces

Friedman tests showed a statistically significant difference between the various interfaces for all sessions – Session 1: $\chi^2(2) = 25.000$, $p < 0.01$, Session 2: $\chi^2(2) = 23.444$, $p < 0.01$, Session 3: $\chi^2(2) = 25.000$, $p < 0.01$, Session 4: $\chi^2(2) = 27.444$, $p < 0.01$, Session 5: $\chi^2(2) = 32.111$, $p < 0.01$. Thus, the null hypothesis $H_{0,1}$ can be rejected for all three interfaces. This result indicates that there is a significant difference in the reload time when using the three interfaces. Table 6.4.1.1 summarises the Wilcoxon post-hoc signed rank test for each session and between each pair of interfaces:

Table 6.4.1.1 - Wilcoxon post-hoc: Reload time

	KM - BCIG	**KM - BCIF**	**BCIG – BCIF**
Session 1	Z = -2.635, p < 0.017*	Z = -3.680, p < 0.017*	Z = -3.593, p < 0.017*
Session 2	Z = -3.201, p < 0.017*	Z = -3.680, p < 0.017*	Z = -3.375, p < 0.017*
Session 3	Z = -3.070, p < 0.017*	Z = -3.593, p < 0.017*	Z = -3.375, p < 0.017*
Session 4	Z = -3.506, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.288, p < 0.017*
Session 5	Z = -3.549, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.680, p < 0.017*

*** = p < 0.017**

As can clearly be seen from Table 6.4.1.1, there was a significant difference between all three interfaces in terms of the mean time taken to reload. Inspection of the mean reload times indicated that KM (6.0 seconds) had significantly faster times than BCIG (8.6 seconds) and BCIF (12.8 seconds). Therefore, KM is better suited for activating a command such as reloading when playing a 2D game and can be recommended for use in this regard. Since this metric measures the time to activate the reload command, the difference between BCIG and BCIF can be solely attributed to the command activation action, namely the Peregrine glove for BCIG and facial expressions for BCIF. That would thus suggest that activating the reload command using facial expressions takes much longer than with the keyboard and glove.

During the user testing it was observed that participants were reluctant to move the cursor while activating a command when making use of BCIF. Moving the cursor required the participants to swivel their head in the direction they wanted to move the cursor, and activating a command required them to produce a certain facial expression. In the case of the reload action, the participants had to raise their eyebrows. It was observed that participants were more comfortable performing the reload action first and then moving their head in order to move the cursor. However, with KM and BCIG the participants were content to move the cursor while they activated the reload action. The reason for this phenomena may be that both the triggers for movement and command activation involved movement of the head and face. This caused a problem for participants as they struggled to perform both actions simultaneously. While this would not have an effect on reload time, it may be further noticed when the initial contact times are analysed.

Sessions

The data was found not to be spherical in all instances (Table 6.4.1.2). The results of the sphericity test and repeated measures ANOVA are tabulated below:

Table 6.4.1.2 - Summary of Mauchley's sphericity test and ANOVA

	KM	BCIG	BCIF
Sphericity test	χ^2 (9) = 44.870	(χ^2 (9) = 46.219*	χ^2 (9) = 67.498*
ANOVA	F (1.818, 30.907) = 28.685*	F (1.702, 28.936) = 37.472*	F (1.342, 22.815) = 12.577*

*** = p < 0.05**

As can be seen from Table 6.4.1.2, $H_{0,2}$ could be rejected for all interfaces, therefore the session has an effect on the reload time. Post-hoc tests were conducted to determine which sessions differed significantly from one another. The results are summarised below (KM is on the topmost row of each cell, BCIG directly below and BCIF on the last row of each cell):

Table 6.4.1.3 - Summary of post-hoc tests: Reload time

	Session 1	Session 2	Session 3	Session 4
Session 2	t(17) = -4.975* t(17) = -7.889* t(17) = -3.151			
Session 3	t(17) = -5.208* t(17) = -6.465* t(17) = -3.865*	t(17) = -1.462 t(17) = -0.495 t(17) = -1.491		
Session 4	t(17) = -6.041* t(17) = -6.008* t(17) = -3.595*	t(17) = -3.794* t(17) = -1.886 t(17) = -1.371	t(17) = -3.453* t(17) = -2.355 t(17) = -0.254	

Session 5	t(17) = -6.581* t(17) = -8.320* t(17) = -4.234*	t(17) = -4.982* t(17) = -5.589* t(17) = -2.732	t(17) = -5.037* t(17) = -4.710* t(17) = -1.882	t(17) = -3.219 t(17) = -1.975 t(17) = -1.868

*** = p < 0.05**

As can clearly be seen from Table 6.4.1.3, the most significant differences can be noticed between session 1 and the rest of the sessions, with the exception of session 2 for BCIF. When inspecting Chart 6.4.1.1, a fast decline in the mean time to reload from session 1 to 2 can be observed.

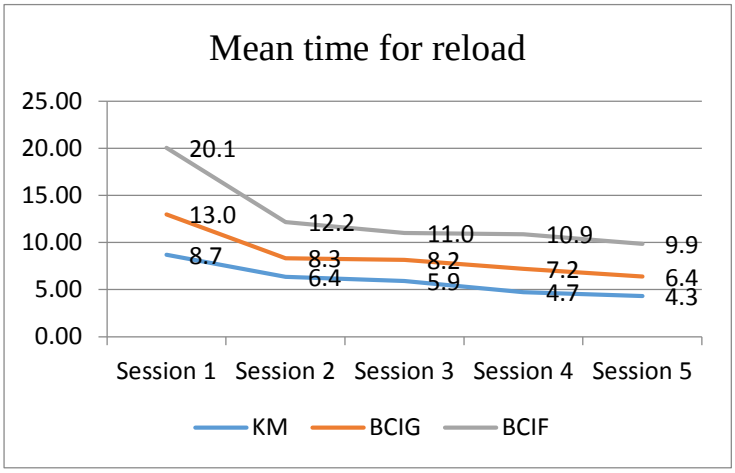

Chart 6.4.1.1 - Mean times for reload.

After this sharp decrease there was a gradual improvement from sessions 2 to 5. This observation was present in all three interfaces, therefore repeated use of the interfaces yielded improved reload times. The sharp decline between sessions 1 and 2 may be attributed to the fact that in the first session participants would forget that the reload first had to be executed before the shoot command would be accepted as an executable command. They would thus immediately attempt to shoot a target when it appeared instead of first reloading in order to shoot. From the second session onwards, the participants remembered to first reload before trying to shoot a target, thus resulting in a faster reload time. This indicates that there was an element of learning in the game itself, but the gradual improvement from sessions 2 to 5 can also be attributed to improved use of the interfaces. Additional sessions might indicate if the steady rate of improvement will continue.

With KM and BCIG, where the triggers for movement and command activation were not situated in close proximity, the participants moved and reloaded at the same time. This behaviour, however, was not present in BCIF where participants first completed an action before moving the cursor. This implies that if this behaviour continued in gameplay, participants would not be able to actively track a moving target due to their inability to move and shoot at the same time. If this is the case, the time on target metric will highlight this

behaviour. This could also result in the initial contact times (section 6.5) of these interfaces being faster than with BCIF, where there was a delay before movement of the cursor was initialised. This phenomena decreased over several sessions as participants grew more accustomed to BCIF while acquiring the skill and confidence to perform the move and reload actions simultaneously. Therefore, use of BCIF improves over time and may be suitable for use in 2D gaming, but requires an element of learning, which was not required with BCIG.

6.4.2. Time to reload summary

KM proved to be the best interface to use when activating a command such as reload for all types of targets in a 2D game. This could be directly attributable to the familiarity of the participants with this interface. BCIG presented the second best reload times for all three tasks and BCIF once again proved to be the slowest of the three interfaces. It was also concluded that significant learning took place during all three tasks, which included learning in terms of the game itself but also and more importantly with the new interfaces.

BCIF was responsible for the best improvement times, this could be explained by the fact that the participants had to activate commands using facial expressions. This interaction was not as natural or intuitive as originally envisioned, thus starting off slowly but with more use they showed better improvement times than with BCIG. BCIG required participants to use their fingers, which emulates the real life act of reloading a shotgun, to activate a command. This action may have been more intuitive to them, thus starting off with a better reload time than with BCIF.

6.5. Initial contact with target

The initial contact time was measured from when the target was presented until the cursor first made contact with the target. This does not include the time that the cursor moved off and back onto the target, which is rather regarded as an overshoot, and will be discussed in Section 6.6.

The following hypotheses were formulated for the analysis of the three tasks:

- $H_{0,1}$: The interface used has no effect on the time taken to make initial contact with the target.
- $H_{0,2}$: There is no difference in the time taken to make initial contact with the target between multiple sessions.

6.5.1. Task 1: Initial contact with the target

In this instance, the assumption of sphericity (χ^2 (9) = 12.878, p > 0.05) was met at an α-level of 0.05. A within-subjects repeated measures ANOVA indicated that there was significant interaction between the two factors, namely session and interfaces (F(8, 204) = 2.265, p < 0.05), therefore separate ANOVAS and Friedman tests were conducted.

Interfaces

Friedman tests showed a statistically significant difference between the different interfaces for all sessions – Session 1: $\chi^2(2) = 27.444$, $p < 0.01$, Session 2: $\chi^2(2) = 31.000$, $p < 0.01$, Session 3: $\chi^2(2) = 30.789$, $p < 0.01$, Session 4: $\chi^2(2) = 31.000$, $p < 0.01$, Session 5: $\chi^2(2) = 30.333$, $p < 0.01$. Therefore, the null hypothesis $H_{0,1}$ can be rejected for all three interfaces. This result indicates that there is a significant difference in the first contact time when using the three user interfaces. Table 6.5.1.1 summarises the Wilcoxon signed rank test for each session and between each pair of interfaces.

Table 6.5.1.1 – Results of Wilcoxon tests: Initial contact times

	KM – BCIG	**KM – BCIF**	**BCIG - BCIF**
Session 1	Z = -3.680, p < 0.017*	Z = -3.680, p < 0.017*	Z = -3.070, p < 0.017*
Session 2	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.375, p < 0.017*
Session 3	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*	Z = -2.675, p < 0.017*
Session 4	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.462, p < 0.017*
Session 5	Z = -3.636, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.506, p < 0.017*

*** p < 0.017**

Inspection of the mean times to make initial contact with the target during task 1 indicated that KM (15.4 seconds) had significantly faster times than BCIG (23.8 seconds) as well as BCIF (29.6 seconds). Based on this, together with the fact that post-tests indicate there is a significant difference between all interfaces, it can be concluded that moving the cursor to make initial contact with the target while making use of KM is significantly faster than using BCIG or BCIF. Therefore, KM is better suited for moving a cursor to a target location and would be the best interface option in this regard.

Although KM is better suited to the specific task, a significant difference in the mean times between BCIG and BCIF was noted. When having the choice between these two interfaces BCIG would be the more suitable interface when moving the cursor to make initial contact with the target. This difference in initial contact times, despite making use of the same method

for cursor control, may be the result of the difficulty participants encountered while attempting to simultaneously move and reload with BCIF, as discussed in Section 6.3.1.

Sessions

The data was found to be spherical in all instances (Table 6.5.1.2). The results of the repeated measures ANOVA are tabulated below:

Table 6.5.1.2 - Results of Mauchley's sphericity test and ANOVA

	KM	BCIG	BCIF
Sphericity test	χ^2 (9) = 7.245	χ^2 (9) = 15.322	χ^2 (9) = 13.020
ANOVA	F (4, 68) = 7.481*	F (4, 68) = 40.827*	F (4, 68) = 23.746*

*** p < 0.05**

As can clearly be seen from Table 6.5.1.2, $H_{0,2}$ could be rejected for all interfaces, therefore the session has an effect on the initial contact time. Since there were significant differences present, post-hoc tests were conducted to determine which sessions differed significantly from the other. The results of the post-hoc tests are summarised below (KM is on the topmost row of each cell, BCIG directly below and BCIF on the last row of each cell):

Table 6.5.1.3 - Post-hoc test results: Initial contact - Task 1

	Session 1	Session 2	Session 3	Session 4
Session 2	t(17) = 2.637 t(17) = 5.325* t(17) = 6.441*			
Session 3	t(17) = 5.601* t(17) = 6.068* t(17) = 5.705*	t(17) = 2.048 t(17) = 1.165 t(17) = 1.648		
Session 4	t(17) = 3.699* t(17) = 9.225* t(17) = 5.773*	t(17) = 1.533 t(17) = 5.069* t(17) = 2.238	t(17) = -0.281 t(17) = 4.298* t(17) = 1.315	
Session 5	t(17) = 4.432* t(17) = 10.148* t(17) = 7.225*	t(17) = 2.378 t(17) = 9.254* t(17) = 4.49*	t(17) = 0.646 t(17) = 4.566* t(17) = 3.608*	t(17) = 1.097 t(17) = 2.403 t(17) = 2.18

*** p < 0.05**

As can clearly be seen from Table 6.5.1.3, the most significant differences are between session 1 and the rest of the sessions, with the exception of session 2 for KM. Thereafter, a steady decline in the time it took to make the initial contact with the target for the three interfaces from session 2 to 5 is evident (Chart 6.5.1.1).

Inspection of the mean times indicated that the initial contact time during session 5 was the fastest for KM, BCIG and BCIF. By inspecting the difference between the means for session 1

and 5 for the three interfaces there was a decline in the time to make initial contact for all interfaces, although small for KM. This indicates that learning did take place for all three interfaces, although the best improvement was noticed for BCIF. As mentioned before, part of the improvement was due to the participants becoming more familiar with the game since improvement was present when using the keyboard as well.

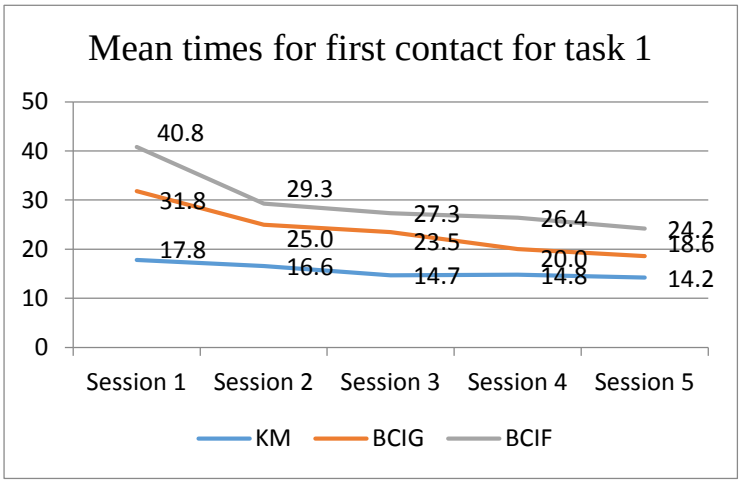

Chart 6.5.1.1 - Mean times for task 1

However, improvement with the other interfaces from sessions 1 to 5 (BCIG - 13.2, BCIF – 16.6 seconds) was more noticeable than with KM (3.6 seconds), thus while there was learning of the game, there is also an element of improvement that can be attributed to BCIG and BCIF. Therefore, it could be established that repeated use of the interfaces yielded improved times for the initial contact metric for all three interfaces.

Based on these findings, it can be argued that the new interfaces can be used for 2D games, although it may take some time for them to reach their potential. The participants included in the study were computer literate and accustomed to making use of the keyboard and mouse. It can therefore be argued that this may influence the acceptance and effective use of the new interfaces. Additionally, most of the participants had not used other devices for computer input before. Prior knowledge of the keyboard and mouse combination is a real world problem that is faced. In order for individuals to migrate from traditional interfaces to more natural interfaces, users will have to be convinced that these NUIs are more usable than the traditional interfaces. It would be interesting to again conduct the study but only include individuals with no computer training which would give a clearer indication of the rate of learning for the three interfaces and acceptance of these three interfaces.

6.5.2. Task 2: Initial contact with the target

In this instance, the assumption of sphericity (χ^2 (9) = 68.979, $p < 0.05$) was not met at an α-level of 0.05. A within-subjects repeated measures ANOVA indicated that there was significant interaction between the two factors, namely session and the interfaces (F(4.332, 110.461) = 3.456, $p < 0.05$), therefore separate ANOVA and Friedman analyses had to be conducted by isolating each factor in turn.

Interfaces

Friedman tests showed a statistically significant difference between the different interfaces for all sessions – Session 1: $\chi^2(2) = 27.444$, $p < 0.01$, Session 2: $\chi^2(2) = 34.111$, $p < 0.01$, Session 3: $\chi^2(2) = 28.141$, $p < 0.01$, Session 4: $\chi^2(2) = 32.444$, $p < 0.01$, Session 5: $\chi^2(2) = 34.111$, $p < 0.01$. Similar to task 1 the null hypothesis $H_{0,1}$ can be rejected for all sessions, therefore the interface used has an effect on the time taken to make initial contact with the target. The table below summarises the Wilcoxon signed rank test for each session and between each pair of interfaces:

Table 6.5.2.1 – Wilcoxon post-hoc: First contact time

	KM - BCIG	**KM – BCIF**	**BCIG - BCIF**
Session 1	Z = -3.637, p < 0.017*	Z = -3.724, p < 0.017*	Z = -2.417, p < 0.017*
Session 2	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.680, p < 0.017*
Session 3	Z = -3.527, p < 0.017*	Z = -3.724, p < 0.017*	Z = -2.940, p < 0.017*
Session 4	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.245, p < 0.017*
Session 5	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.680, p < 0.017*

*** = p < 0.017**

As can clearly be seen from Table 6.5.2.1, there was a significant difference between all three interfaces in terms of the mean time taken to make initial contact with the target during task 2. As with the previous metrics, the mean time over 5 sessions indicated that KM was significantly faster (16.4 seconds) than both BCIG (24.1 seconds) and BCIF (30.8 seconds). This is quite possibly due to the familiarity of the interface, thus indicating that BCIG and BCIF require some learning. The difference between BCIG and BCIF can once again be explained by the observation that was made where participants first completed the reload action when using the facial expressions before moving the cursor, thus increasing the time to make contact with the target.

Sessions

The data was found to be not spherical in all instances (Table 6.5.2.2). The results of the repeated measures ANOVA are tabulated below:

Table 6.5.2.2 - Results of Mauchley's sphericity test and ANOVA

	KM	BCIG	BCIF
Sphericity test	χ^2 (9) = 24.292*	χ^2 (9) = 27.697*	χ^2 (9) = 29.373*
ANOVA	F (2.111, 35.882) = 8.935*	F (2.665, 45.310) = 27.479*	F (1.936, 32.911) = 11.092*

*** p < 0.05**

As can clearly be seen from Table 6.5.2.2, $H_{0,2}$ could be rejected for all interfaces, therefore the session has an effect on the initial contact time. Since there were significant differences, post-hoc tests were conducted. The results of the post-hoc tests are summarised below (KM is on the topmost row of each cell, BCIG directly below and BCIF on the last row of each cell):

Table 6.5.2.3 - Post hoc results: First contact time - Task 2

	Session 1	Session 2	Session 3	Session 4
Session 2	t(17) = -3.44* t(17) = -5.055* t(17) = -1.534			
Session 3	t(17) = -2.963 t(17) = -6.523* t(17) = -3.569*	t(17) = - 0.165 t(17) = -1.706 t(17) = -4.305*		
Session 4	t(17) = -3.434* t(17) = -4.801* t(17) = -2.448	t(17) = -1.219 t(17) = -1.171 t(17) = -2.32	t(17) = -0.834 t(17) = 0.2 t(17) = 1.463	
Session 5	t(17) = -3.862* t(17) = -8.693* t(17) = -4.789*	t(17) = -2.412 t(17) = -10.404* t(17) = -5.59*	t(17) = -2.046 t(17) = -4.678* t(17) = -1.727	t(17) = -1.066 t(17) = -4.014* t(17) = -3.095

*** p < 0.05**

As seen in Table 6.5.2.3, there were significant differences between some of the sessions. With the exception of session 4, there was a steady decline for the new interfaces in terms of the time it took to make the initial contact, as seen in Chart 6.5.2.1. Inspection of the mean times indicated that the initial contact time during session 5 was the fastest for all three interfaces, indicating that learning did take place.

The best improvement was noticed for BCIF and this indicates that the facial expressions became easier to handle over time. Making use of facial expressions as input was a very unfamiliar action for the participants. During the first session the participants experienced some difficulty using BCIF, conversely the use of the Peregrine glove that makes use of finger movements, was more intuitive to them, thus leading to better initial contact times during session 1 for BCIG. This resulted in a faster starting time for BCIG and left less room for

improvement. It may thus be argued that the fact that the participants were used to the keyboard and mouse where they make use of finger movements for input, may have had an effect on their performance with the new interfaces.

Therefore, it could be established that repeated use of the interfaces yielded improved times for the initial contact time metric.

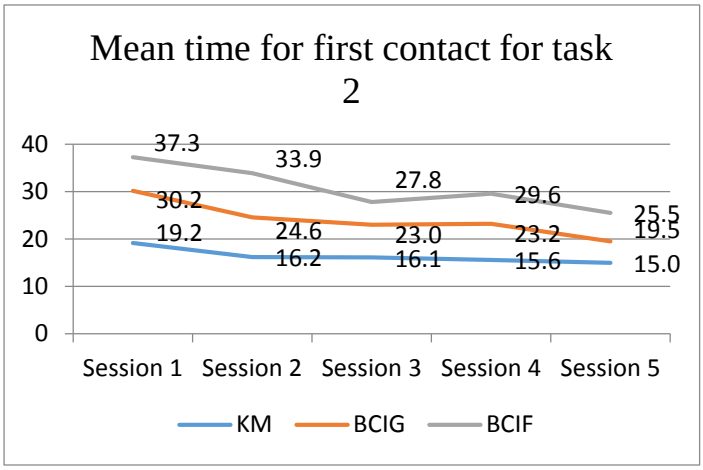

Chart 6.5.2.1 - Mean times for task 2

6.5.3. Task 3: Initial contact with the target

The assumption of sphericity (χ^2 (9) = 78.378, $p < 0.05$) was not met at an α-level of 0.05. A within-subjects repeated measures ANOVA indicated that there was significant interaction between the two factors, namely session and the interfaces (F (4.135, 105.441) = 4.732, $p < 0.05$), resulting in separate ANOVAS and Friedman tests being conducted.

Interfaces

Friedman tests showed a statistically significant difference between the different interfaces for all sessions – Session 1: $\chi^2(2) = 26.333$, $p < 0.01$, Session 2: $\chi^2(2) = 27.086$, $p < 0.01$, Session 3: $\chi^2(2) = 27.086$, $p < 0.01$, Session 4: $\chi^2(2) = 32.444$, $p < 0.01$, Session 5: $\chi^2(2) = 34.111$, $p < 0.01$.

Thus, the null hypothesis $H_{0,1}$ can be rejected for all sessions. Therefore the interface used has an effect on the time taken to make initial contact with the target, when facing unpredictable moving targets. This result corresponds with the result for tasks 1 and 2.

The table below summarises the Wilcoxon signed rank test for each session and between each pair of interfaces:

Table 6.5.3.1 - Post-hoc tests: Initial contact times

	KM – BCIG	KM - BCIF	BCIG - BCIF
Session 1	Z = -3.680, p < 0.017*	Z = -3.724, p < 0.017*	Z = -1.851, p > 0.017
Session 2	Z = -3.516, p < 0.017*	Z = -3.724, p < 0.017*	Z = -2.940, p < 0.017*
Session 3	Z = -3.516, p < 0.017*	Z = -3.724, p < 0.017*	Z = -2.940, p < 0.017*
Session 4	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.288, p < 0.017*
Session 5	Z = -3.680, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*

*** = p < 0.017**

As can clearly be seen from Table 6.5.3.1, there was a significant difference between all three interfaces in terms of the mean time taken to make initial contact with the target during task 2, with the exception of BCIG and BCIF for session 1. Inspection of the mean times to complete task 1 indicated that KM (17.5 seconds) had significantly faster times than BCIG (27.0 seconds) as well as BCIF (33.1 seconds). Therefore, it could be concluded that moving the cursor to make initial contact while making use of KM is significantly faster than using BCIG or BCIF. Consequently, KM is better suited for moving the cursor to intercept an unpredictable moving target and can be recommended for use in this regard. This could be attributed to the participants being more accustomed to stabilising the mouse cursor while using the traditional mouse.

A significant difference in the mean times between BCIG and BCIF was also noted. Therefore, when having the choice between these two interfaces, BCIG would be the more suitable interface when moving the cursor to make initial contact with an unpredictable moving target. The difference between BCIG and BCIF can once again be explained by the observation that difficulty was experienced with BCIF when attempting to simultaneously activate a command as well as control the cursor, as discussed in Chapter 6.5.1.

Sessions

The data was found to be spherical for KM but not for BCIG and BCIF (Table 6.5.3.2). The results of the repeated measures ANOVA are tabulated on the next page:

Table 6.5.3.2 - Results of Mauchley's sphericity test and ANOVA

	KM	BCIG	BCIF
Sphericity test	χ^2 (9) = 16.416	χ^2 (9) = 29.577*	χ^2 (9) = 36.974*
ANOVA	F (4, 68) = 10.595*	F (2.239, 38.057) = 20.097*	F (1.769, 30.071) = 17.896*

*** = p < 0.05**

As can clearly be seen from the table, $H_{0,2}$ could be rejected for all interfaces, therefore the session has an effect on the initial contact time. Post-hoc tests were conducted to determine which sessions differed significantly from the other. The results of the post-hoc tests are summarised below (KM is on the topmost row of each cell, BCIG directly below and BCIF on the last row of each cell):

Table 6.5.3.3 - Post-hoc test results: Task 3

	Session 1	**Session 2**	**Session 3**	**Session 4**
Session 2	t(17) = -2.178 t(17) = -4.016* t(17) = -3.278*			
Session 3	t(17) = -4.927* t(17) = -5.139* t(17) = -4.622*	t(17) = -2.482 t(17) = -2.41 t(17) = -2.736		
Session 4	t(17) = -4.054* t(17) = -4.639* t(17) = -4.365*	t(17) = -2.378 t(17) = -1.61 t(17) = -4.206*	t(17) = -0.733 t(17) = 0.018 t(17) = -0.135	
Session 5	t(17) = -4.237* t(17) = -6.876 t(17) = -5.346*	t(17) = -3.227 t(17) = -4.39* t(17) = -6.145*	t(17) = -1.511 t(17) = -2.267 t(17) = -2.054	t(17) = -0.606 t(17) = -2.652 t(17) = -2.560

*** $p < 0.05$**

As can be seen, the most significant differences can generally be noticed between session 1 and the remainder of the sessions. Although the most significant differences were present from session 1 to 2, there was a steady decline in the remainder of the sessions for the new interfaces namely BCIG and BCIF in the time it took to make the initial contact with the target. KM, which the participants were at ease with, showed an improvement from sessions 1 to 2 (Chart 6.5.3.1), where-after the decrease in time levelled out, indicating that there was learning of the game more so than improvement with the interface. However, with BCIG and BCIF, there was continued improvement with the exception of session 4, which can be reasoned, included both improvement with the game as well as with the interfaces.

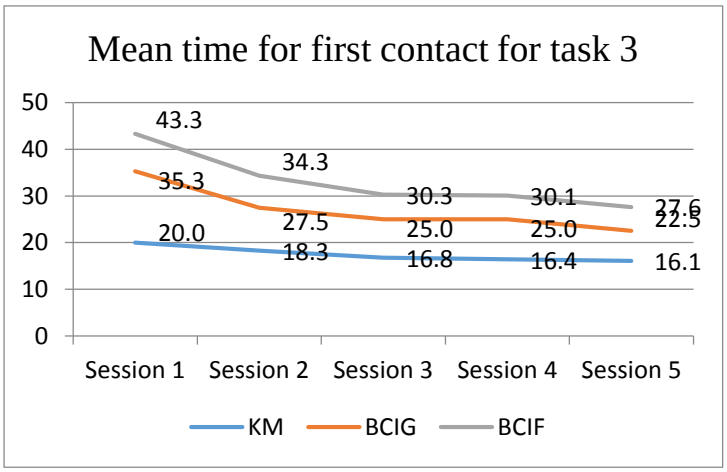

Chart 6.5.3.1 - Mean times for task 3

Inspection of the mean times indicated that the initial contact time during session 5 was the fastest for KM, BCIG and BCIF. Inspection of the difference between the means for sessions 1 and 5 for the three interfaces shows a decline in the time to make initial contact for all interfaces. This indicates that learning took place for all three interfaces, although the best improvement was noticed for BCIF. Therefore, it could be established that repeated use of the interfaces yielded improved times for the initial contact time metric.

BCIG and BCIF made use of the same action for cursor movement, nonetheless a difference in the mean times can be noticed, which corresponds with the results of tasks 1 and 2. This indicates that the action used to activate commands had an influence on the time it took to make initial contact for BCIG and BCIF since the reload action was first completed before repositioning the cursor. Participants were observed struggling to perform both actions simultaneously when using BCIF, as a result the reload action was completed before movement of the cursor was attempted, whereas when using the glove, participants reloaded while moving the cursor to the target. This may explain the difference in first contact times between the mentioned interfaces. This effect could be eliminated by only revealing the target once the reload action has been performed, which was not the case in this study.

6.5.4. Initial contact time summary

KM achieved the fastest initial contact times for all three tasks, thus KM is the best interface to use when attempting to position the cursor over a stationary, predictable and unpredictable target.

Although the BCI was used for cursor movement in both BCIG and BCIF, a difference in the mean times for first contact could be noted between them. This could be due to both cursor movement and activation of commands being conducted using facial and head movement. This can be further supported by inspecting the mean improvement times for BCIG and BCIF. For all tasks, these two interfaces showed dissimilar improvement times. The improvement times for Task 1 was 13.2 seconds for BCIG and 16.6 seconds for BCIF, for Task 2 the times were 10.6 seconds for BCIG and 11.9 seconds for BCIF and finally for Task 3 the times were 12.7 seconds for BCIG and 15.7 seconds for BCIF. This difference in improvement, while making use of the same action for cursor movement, indicates that something other than cursor movement was responsible for the difference in improvement. Consequently it can be reasoned that the method for command activation was responsible for the difference in improvement,

indicating that there was a faster rate of improvement with the use of facial expressions than with the Peregrine glove. It can be argued that the Peregrine glove showed less improvement due to the device initially being more intuitive.

Nonetheless, improvements were noticed for all three interfaces for all tasks, thus indicating that learning took place. Since this metric mainly focused on cursor control, it can be inferred that the actual learning that took place was in terms of the pointing device used. Consequently, with repeated use of the head-mounted accelerometer the participants were able to improve their cursor control and speed.

6.6. Number of overshoots

The number of overshoots was measured by totalling the number of times the participant positioned the cursor over the target and then moved it off the target without eliminating the target.

The following hypotheses were formulated for the analysis of the three tasks:

- $H_{0,1}$: The interface used has no effect on the number of overshoots.
- $H_{0,2}$: There is no difference in the number of overshoots between multiple sessions.

6.6.1. Task 1: Number of overshoots

The assumption of sphericity (χ^2 (9) = 64.650, $p < 0.05$) was not met at an α-level of 0.05. A within-subjects repeated measures ANOVA indicated that there was significant interaction between the two factors, namely session and interfaces (F (4.615, 117.686) = 6.778, $p < 0.05$), therefore separate analyses had to be conducted by isolating each factor in turn.

Interfaces

Friedman tests showed a statistically significant difference between the different interfaces for all sessions – Session 1: $\chi^2(2) = 32.444$, $p < 0.01$, Session 2: $\chi^2(2) = 22.845$, $p < 0.01$, Session 3: $\chi^2(2) = 24.087$, $p < 0.01$, Session 4: $\chi^2(2) = 28.229$, $p < 0.01$, Session 5: $\chi^2(2) = 25.444$, $p < 0.01$. Consequently the null hypothesis $H_{0,1}$ can be rejected for all three interfaces for all sessions. Therefore, the interface used has an effect on the number of overshoots. The table below summarises the Wilcoxon signed rank test for each session and between each pair of interfaces:

Table 6.6.1.1 - Summary of Wilcoxon post-hoc tests: Number of overshoots

	KM - BCIG	KM – BCIF	BCIG – BCIF
Session 1	Z = -3.728, p < 0.017*	Z = -3.725, p < 0.017*	Z = -3.577, p < 0.017*
Session 2	Z = -3.382, p < 0.017*	Z = -3.627, p < 0.017*	Z = -2.858, p < 0.017*
Session 3	Z = -3.534, p < 0.017*	Z = -3.660, p < 0.017*	Z = -2.924, p < 0.017*
Session 4	Z = -3.729, p < 0.017*	Z = -3.727, p < 0.017*	Z = -0.751, p > 0.017
Session 5	Z = -3.661, p < 0.017*	Z = -3.726, p < 0.017*	Z = -1.512, p > 0.017

*** = p < 0.017**

As can clearly be seen from Table 6.6.1.1, there was a significant difference between all three interfaces in terms of the mean number of overshoots during task 1 with the exception of interface 2 and 3 for session 4 and 5. This indicates that although there was a difference in the number of overshoots to start with, the two interfaces did not show a significant difference during the last two sessions. Although both interfaces used the same action for cursor movement, it could be reasoned that BCIG was more intuitive than BCIF. This indicates that repeated use of BCIF was required before BCIG and BCIF could be used at a similar level of accuracy.

Inspection of the means indicated that KM (3.0) had significantly less overshoots than BCIG (9.8) as well as BCIF (15.3). Once again KM is better suited to the task at hand. It also indicates that BCIG would be the more suitable interface when targeting objects if KM is not available.

Although the BCI again was used for mouse movement in both BCIG and BCIF a difference in the means for the number of overshoots was noted. This could be due to both cursor movement and activation of commands being conducted using facial and head movement when using BCIF. During user testing it was observed that when participants attempted to perform the shoot action, while using facial expressions, they tended to move their head slightly. This unintended head movement caused the cursor to move off target and resulted in an overshoot. Thus, the action used for triggering the shoot command could be responsible for the difference.

Sessions

The data was found to be spherical for KM but not for BCIG and BCIF (Table 6.6.1.2). The results of the repeated measures ANOVA are tabulated on the next page:

Table 6.6.1.2 - Summary of Mauchley's sphericity test and ANOVA

	KM	BCIG	BCIF
Sphericity test	χ^2 (9) = 7.859	χ^2 (9) = 17.484*	χ^2 (9) = 42.617*
ANOVA	F (4, 68) = 7.860*	F (2.392, 40.672) = 1.615	F (1.682, 28.590) = 9.782*

*** p < 0.05**

$H_{0,2}$ could be rejected for KM and BCIF, as can clearly be seen from the table, therefore the session has an effect on the number of overshoots for these two interfaces, but not for BCIG. Since there were significant differences, post-hoc tests were conducted. The results of the post-hoc tests are summarised below (KM is on the topmost row of each cell and BCIF directly below):

Table 6.6.1.3 - Summary of post-hoc tests: Number of overshoots

	Session 1	Session 2	Session 3	Session 4
Session 2	t(17) = 0.248 t(17) = -3.904*			
Session 3	t(17) = 0.669 t(17) = -3.067	t(17) = 0.283 t(17) = 0.683		
Session 4	t(17) = -3.609* t(17) = -4.011*	t(17) = -3.030 t(17) = -1.287	t(17) = -5.659* t(17) = -1.821	
Session 5	t(17) = -3.220 t(17) = -3.246*	t(17) = -2.971 t(17) = -0.973	t(17) = -4.306 t(17) = -1.513	t(17) = 0.707 t(17) = 0.292

*** p < 0.05**

KM only showed a significant difference between sessions 1 and 4 as seen in Table 6.6.1.3. There was a small improvement with KM, but BCIF showed a dramatic improvement from session 1 to 2. As can be seen in Table 6.6.1.3, there is a significant difference between session 1 and sessions 2, 4, 5 for BCIF. There was a fast decline in the mean number of overshoots from session 1 to 2 (improvement of 10.7), where after the number of overshoots stabilised and did not fluctuate by more than three overshoots between sessions (Chart 6.6.1.1).

Inspection of Chart 6.6.1.1 revealed that KM averaged 3.7 overshoots during session 1, and 2 overshoots for session 5, an improvement of 1.7 overshoots. Thus, the number of overshoots started low and showed a slight decrease during the later sessions, which could be expected due to the frequent use of this interface.

Once again, if analysed in context, the means in Chart 6.6.1.1 do not represent large improvements or declines. The number of overshoots for BCIG was stable through the 5 sessions, not fluctuating by more than 2.3 overshoots from one session to the next.

As seen in Table 6.6.1.2 and Chart 6.6.1.1, it is clear that there was a significant improvement from the first session to the later sessions for KM and BCIF. A small improvement was noted for BCIG. The improvement with BCIF from session 1 to 5 was more noticeable than with KM and BCIG, thus while there was learning of the game, there is also an element of improvement that can be attributed to BCIF itself.

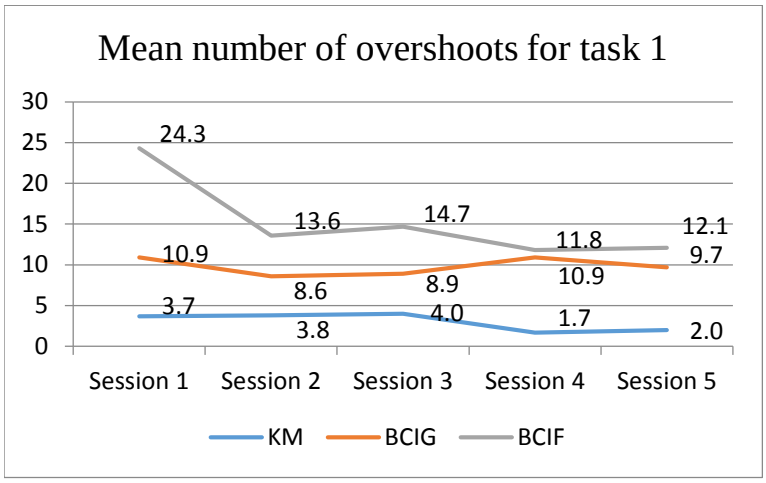

Chart 6.6.1.1 - Mean number of overshoots

The fact that BCIG did not show a significant difference between the different sessions may show that the glove used in combination with the BCI lead to more stable positioning than the facial expressions. BCIG and KM showed similar improvements from session 1 to 5. While BCIG was responsible for more stable cursor control than BCIF, after 5 sessions it could not achieve a comparable number of overshoots to KM. Consequently, it cannot compare to the effectiveness of KM for 2D gameplay. Problems that were encountered with the facial expressions as discussed earlier may be the cause of the large number of overshoots for BCIF, as the use of this particular facial expressions caused small movements of the head and thus the cursor. In this study winking was used to activate the shoot command, and caused the participants to move their heads in the direction of the eye wink (left when performing a left wink and right when performing a right wink). Therefore, using another facial expression that does not cause unwanted movement of the head could eliminate this problem.

Therefore, it could be established that repeated use of the interfaces yielded improvement in the number of overshoots for two of the interfaces.

6.6.2. Task 2: Number of overshoots

The assumption of sphericity (χ^2 (9) = 44.695, $p < 0.05$) was not met at an α-level of 0.05 for task 2. A within-subjects repeated measures ANOVA indicated that there was significant interaction between the two factors (F (5.590, 142.554) = 15.627, $p < 0.05$), therefore separate ANOVAS and Friedman tests were conducted.

Interfaces

Friedman tests showed a statistically significant difference between the different interfaces for all sessions – Session 1: $\chi^2(2) = 24.333$, $p < 0.01$, Session 2: $\chi^2(2) = 27.444$, $p < 0.01$, Session 3: $\chi^2(2) = 26.225$, $p < 0.01$, Session 4: $\chi^2(2) = 25.800$, $p < 0.01$, Session 5: $\chi^2(2) = 18.778$, $p < 0.01$, leading to the rejection of the null hypothesis $H_{0,1}$ for all three interfaces. This result indicates that there is a significant difference in the number of overshoots when using the three user interfaces. The table below summarises the Wilcoxon signed rank test for each session and between each pair of interfaces:

Table 6.6.2.1 - Summary of Wilcoxon post-hoc tests: Number of overshoots

	KM - BCIG	**KM - BCIF**	**BCIG - BCIF**
Session 1	Z = -1.264, p > 0.017	Z = -3.681, p < 0.017*	Z = -3.724, p < 0.017*
Session 2	Z = -1.264, p > 0.017	Z = -3.681, p < 0.017*	Z = -3.724, p < 0.017*
Session 3	Z = -2.156, p > 0.017	Z = -3.724, p < 0.017*	Z = -3.682, p < 0.017*
Session 4	Z = -3.272, p < 0.017*	Z = -3.680, p < 0.017*	Z = -3.432, p < 0.017*
Session 5	Z = -2.527, p < 0.017*	Z = -3.681, p < 0.017*	Z = -2.704, p < 0.017*

*** = p < 0.017**

As can clearly be seen from Table 6.6.2.1, there was a significant difference between all three interfaces in terms of the mean number of overshoots during task 2, with the exception of interface 1 and 2 for sessions 1, 2 and 3. Inspection of the means indicated that KM (7.4) had significantly less overshoots than BCIG (11.5) and BCIF (38). Therefore, it could be concluded that positioning the cursor over the target while making use of KM is significantly more effective than using BCIG or BCIF. If the choice had to be made, BCIG would, in the absence of KM, be the best alternative to use.

As previously mentioned, this could be due to both cursor movement and activation of commands being facilitated using facial and head movement when using BCIF. The difference is more prominent during task 2, where the difference in the mean number of overshoots between BCIG and BCIF is 26.5. During task 1 the mean difference was 5.5. Thus, the fact that

the target did not remain stationary had an impact on the number of overshoots. This further indicates that the participants experienced difficulty in activating the shoot command while making use of BCIF. This led to the target moving away from the cursor as the shoot command was not triggered in time, resulting in an additional overshoot. The observation was also made that when participants attempted to activate the shoot command by using facial expressions, it resulted in unintended head movement that moved the cursor away from the target. This small movement away from the target further compounded the difficulty that the participants experienced with making small adjustments to again position the crosshair over the target. The shoot command may have been activated with the cursor no longer positioned over the target, resulting in a miss, which can be confirmed by a large number of misses for BCIF.

Additionally, through observation it was noted that during earlier sessions participants moved and placed the cursor in the path of the moving target and then attempted to activate the shoot command as the target moved over the cursor. Therefore, the tendency of the participants to ambush the target instead of actively tracking it is indicative of the inclination to compartmentalise actions and movements when using the head for both, which led to more overshoots during the earlier sessions.

Sessions

The data was found to be spherical for BCIG but not for KM and BCIF. These results and those of the repeated measures ANOVA are tabulated below:

Table 6.6.2.2 - Summary of Mauchley's sphericity test and ANOVA

	KM	**BCIG**	**BCIF**
Sphericity test	χ^2 (9) = 18.067*	χ^2 (9) = 9.249	χ^2 (9) = 19.071*
ANOVA	F (2.877, 48.904) = 4.612*	F (4, 68) = 0.736	F (2.592, 44.070) = 20.972*

* $p < 0.05$

As can clearly be seen from Table 6.6.2.2, $H_{0,2}$ could be rejected for KM and BCIF, but not for BCIG, therefore the session has an effect on the number of overshoots for these two interfaces. The results of the post-hoc tests are summarised in Table 6.6.2.3 (KM is on the topmost row of each cell and BCIF directly below):

Table 6.6.2.3 - Summary of post-hoc tests: Number of overshoots

	Session 1	**Session 2**	**Session 3**	**Session 4**
Session 2	t(17) = -0.145 t(17) = -3.361*			

Session 3	t(17) = -0.761 t(17) = -5.719*	t(17) = -0.987 t(17) = -0.922		
Session 4	t(17) = -2.903 t(17) = -5.265*	t(17) = -3.916* t(17) = -1.791	t(17) = -3.052 t(17) = -1.353	
Session 5	t(17) = -2.143 t(17) = -7.358*	t(17) = -3.811 t(17) = -4.537*	t(17) = -1.837 t(17) = -4.534*	t(17) = 0.597 t(17) = -3.621*

*** p < 0.05**

As can be seen in the table above, there is a significant difference in some of the sessions. When inspecting the mean number of overshoots it becomes clear that BCIF showed a large improvement, while KM showed a steady decrease and BCIG showed little improvement. Once again it could be reasoned that BCIG was more intuitive than BCIF, although both used the same action for cursor movement. This indicates that learning first had to occur with BCIF before it could be effectively used to position the cursor and eliminate the target.

When inspecting the means (Chart 6.6.2.1) for KM it is evident that during session 1 a mean of 9.4 overshoots occurred. During sessions 2 and 3 a small decrease was noted, but during sessions 4 and 5 it decreased even more. Due to the non-stationary nature of the targets it appears as though the participants first recognised the predictability of the targets over the first three sessions, while slowly improving their targeting.

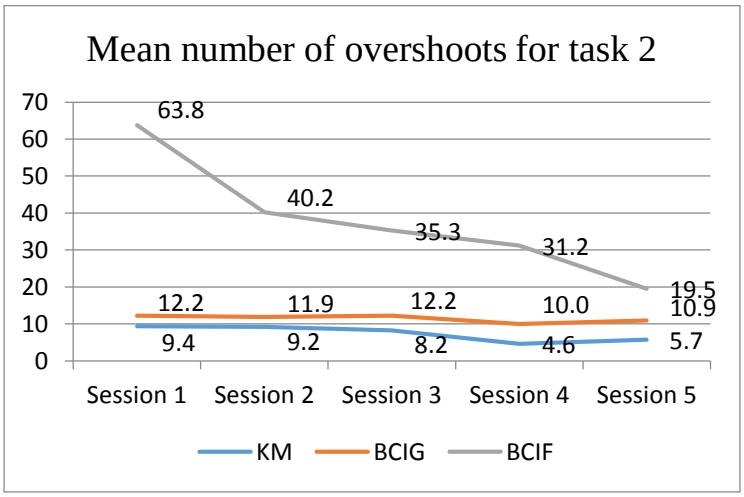

Chart 6.6.2.1 - Mean number of overshoot

Since this interface is frequently used by the participants it can be reasoned that they first had to learn the game before they could apply their competency with KM to the specific task. It can again be reasoned that BCIG is more intuitive than BCIF, as BCIG started out with 0.8 overshoots per target and continued with similar performances during the remaining sessions, where BCIF started with 4.3 overshoots per target and improved to 1.3 overshoots per target during session 5. Thus, extensive learning took place for BCIF. It would be interesting to

investigate whether the decline in the number of overshoots would have continued with further sessions. This indicates that BCIF was not very intuitive to use, but with some training the participants improved and were able to make effective use of the interface.

During sessions 4 and 5 larger improvements were noted when compared to the first three sessions, indicating that participants were able to predict where the target would appear and in which direction it would move. The participants used this knowledge to pre-emptively start to move the cursor to where the target would appear, therefore they were able to position the cursor in the path of the target. This eliminated the need for active tracking, resulting in a more stable position to activate the shoot command from. This may have led to the decrease in the number of overshoots as stated.

This again leads to the idea of effective use of this interface by disabled users. Although it might not be very intuitive, with training it could pose to be a very effective interface for users with little to no mobility below the neck.

6.6.3. Task 3: Number of overshoots

The assumption of sphericity (χ^2 (9) = 46.143, $p < 0.05$) was not met at an α-level of 0.05 for task 3. A within-subjects repeated measures ANOVA indicated that there was significant interaction between the sessions and interfaces ($F(5.603, 142.880) = 8.473$, $p < 0.05$), therefore separate analyses had to be conducted.

Interfaces

Friedman tests showed a statistically significant difference between the different interfaces for all sessions – Session 1: $\chi^2(2) = 28.778$, $p < 0.01$, Session 2: $\chi^2(2) = 23.739$, $p < 0.01$, Session 3: $\chi^2(2) = 28.761$, $p < 0.01$, Session 4: $\chi^2(2) = 25.400$, $p < 0.01$, Session 5: $\chi^2(2) = 16.435$, $p < 0.01$. Thus, the interfaces showed significant difference from each other for session 1 to 5, consequently the null hypothesis $H_{0,1}$ could be rejected for all three interfaces. This result indicates that there is a significant difference in the number of overshoots when using the three user interfaces when faced with unpredictable moving targets. Table 6.6.3.1 summarises the Wilcoxon signed rank test for each session and between each pair of interfaces:

Table 6.6.3.1 - Summary of Wilcoxon post-hoc tests: Number of overshoots

	KM - BCIG	KM - BCIF	BCIG - BCIF
Session 1	Z = -2.618, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*
Session 2	Z = -1.346, p > 0.017	Z = -3.681, p < 0.017*	Z = -3.622, p < 0.017*
Session 3	Z = -1.114, p > 0.017	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*
Session 4	Z = -2.536, p < 0.017*	Z = -3.680, p < 0.017*	Z = -3.624, p < 0.017*
Session 5	Z = -1.996, p > 0.017	Z = -3.681, p < 0.017*	Z = -3.293, p < 0.017*

*** = p < 0.017**

Once again there was a significant difference between all three interfaces, with the exception being between KM and BCIG for sessions 2 and 3. Inspection of the means revealed that KM (8.6 overshoots – over 5 sessions) had significantly less overshoots than BCIG (12.0 overshoots – over 5 sessions) and BCIF (37.6 overshoots – over 5 sessions). This indicates that by using KM, far less overshoots will occur resulting in faster and more effective elimination of unpredictable moving targets. There is a large difference in the mean number of overshoots for BCIG and BCIF, indicating that BCIG would be the better option if a choice had to be made between the two interfaces.

Although BCIG and BCIF made use of the same action for cursor movement, a difference in the mean times can be noticed, similar to the results of tasks 1 and 2. This indicates that the action used to activate commands had an influence on the number of overshoots for BCIG and BCIF. As was discussed during the reload metric, the observed behaviour of participants when using facial expressions can explain this phenomena.

Sessions

The data was found to be spherical for BCIG and BCIF but not for KM (Table 6.6.3.2). The results of the repeated measures ANOVA are tabulated below:

Table 6.6.3.2 - Summary of Mauchley's sphericity test and ANOVA

	KM	BCIG	BCIF
Sphericity test	χ^2 (9) = 17.267*	χ^2 (9) = 15.127	χ^2 (9) = 16.687
ANOVA	F (2.703, 45.951) = 6.557*	F (4, 68) = 6.023*	F (4, 68) = 17.243*

*** p < 0.05**

As can clearly be seen from Table 6.3.3.2, $H_{0,2}$ could be rejected for all interfaces, therefore the session has an effect on the number of overshoots. Since there were significant differences, post-hoc tests were conducted to determine which sessions differed significantly from the

other. The results of the post-hoc tests are summarised below (KM is on the topmost row of each cell, BCIG directly below and BCIF on the last row of each cell):

Table 6.6.3.3 - Summary of post-hoc tests: Number of overshoots

	Session 1	**Session 2**	**Session 3**	**Session 4**
Session 2	t(17) = -2.253 t(17) = -2.482 t(17) = -2.534			
Session 3	t(17) = -2.330 t(17) = -2.903 t(17) = -3.567*	t(17) = -0.390 t(17) = -0.856 t(17) = -2.42		
Session 4	t(17) = -3.437* t(17) = -3.539* t(17) = -3.957*	t(17) = -2.065 t(17) = -1.26 t(17) = -2.182	t(17) = -2.183 t(17) = -0.844 t(17) = -0.354	
Session 5	t(17) = -4.553* t(17) = -3.31* t(17) = -7.479*	t(17) = -1.618 t(17) = -0.919 t(17) = -6.918*	t(17) = -1.728 t(17) = -0.182 t(17) = -4.265*	t(17) = 0.609 t(17) = 0.79 t(17) = -4.277*

*** p < 0.05**

As can be seen in the table above, there is a significant difference in some of the sessions. Although the BCI again was used for cursor movement in both BCIG and BCIF a difference in the means for the number of overshoots was noted, as seen in Chart 6.6.3.1.

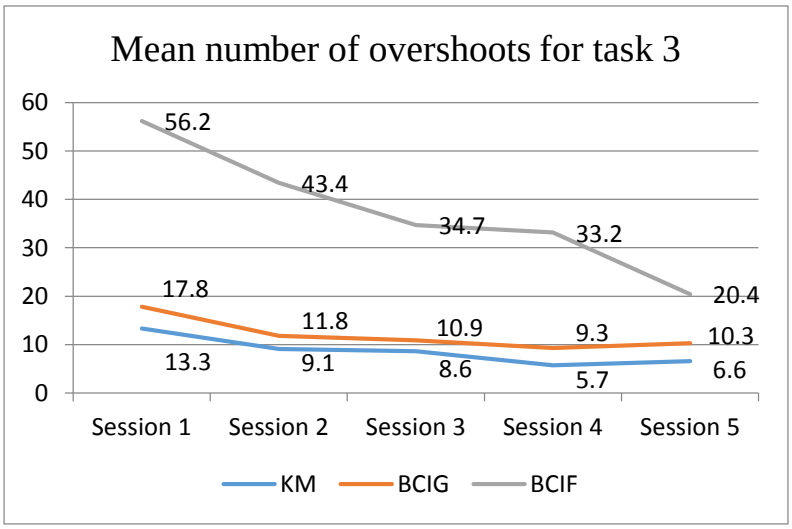

Chart 6.6.3.1 - Mean number of overshoots for task 3

There was a steady decline for the new interfaces, namely BCIG and BCIF, with the exception of session 5 for BCIG, as seen in Chart 6.6.3.1. During task 3, all three interfaces showed improvement, therefore, it could be established that repeated use of the interfaces yielded improved times for the initial contact time metric, therefore learning took place.

6.6.4. Number of overshoots summary

The analysis of the three tasks indicated that KM had the least amount of overshoots and was thus the best interface to use in this regard. Although the BCI was used for mouse movement in both BCIG and BCIF a difference in the means for the number of overshoots could be noted. Through observation during user testing it was noticed that participants, when using BCIF, had trouble keeping the cursor positioned over the target while it was moving. It was observed that most participants' heads moved when winking and thus moved the cursor off target before the shot could be fired, thus resulting in an overshoot. This observation confirms that of Logsdon (2011), where the same part of the body (the users' hand) was used for command activation as well as cursor control, resulting in unintentional movement of the hand, which led to accuracy problems (Logsdon, 2011).

When comparing task 1 with tasks 2 and 3 it becomes clear that a moving target has an influence on the number of overshoots for BCIF. During task 1, BCIG had a mean number of overshoots of 9.8, whereas BCIF had a mean number of 15.3. During task 2 this changed drastically to a mean of 11.5 for BCIG and 38 for BCIF. This same phenomenon is seen in task 3 where BCIG had a mean of 12 compare to BCIF with a mean of 37.6.

This clearly indicates that participants had less trouble positioning the cursor over a static target than with a moving target. This in turn could explain why the number of overshoots for BCIF were more than that of BCIG although both of them used the same action for cursor movement. Thus, the action used for triggering the shoot command could be responsible for the difference since the same method was used for cursor control. The difficulty experienced with the activation of the shoot command when using a right or left wink, where the command activation was accompanied by a small head movement, may be further compounded in the case of moving targets, since the cursor and the target are both moving during the participant's attempt to eliminate the target. When the target is stationary, only the cursor moves, so it is possible that after the movement the cursor is still on the target. Consequently there were more overshoots when using facial expressions than when using the Peregrine glove.

However, the improvement with BCIF for all three tasks was more noticeable than with KM and BCIG, therefore while there was learning of the game, there was also an element of improvement that can be attributed to BCIF itself. As participants grew more accustomed to the interface the tracking of a moving target drastically improved, indicating that BCIF was not

very intuitive to use at first. The fact that BCIG showed no significant improvement during tasks 1 and 2 may show that the glove used in combination with the BCI mouse may be more intuitive to use than the facial expressions used in BCIF. Therefore participants were quicker to use the device to its full potential and did not need to spend time becoming accustomed with the device.

6.7. Time on target

The time on target metric was measured from when the cursor was positioned over the target until the target was eliminated, provided that the cursor did not leave the target during this time. The following hypotheses were formulated for the analysis:

- $H_{0,1}$: The interface used has no effect on the time on target since the last overshoot.
- $H_{0,2}$: There is no difference in the time on target between multiple sessions.

6.7.1. Task 1: Time on target

In this instance, the assumption of sphericity (χ^2 (9) = 16.769, $p > 0.05$) was met at an α-level of 0.05. A repeated measures within-subjects ANOVA was performed to analyse the afore-mentioned hypotheses. Post-hoc tests were applied to determine which interfaces or sessions were responsible for the significant difference in the time on target since the last overshoot when significant differences were detected with the ANOVA.

Interfaces

At an α-value of 0.05 the null hypothesis $H_{0,1}$ can be rejected (F (2, 51) = 38.206, $p < 0.05$), therefore the interface used has an effect on the time spent on target since the last overshoot. As can clearly be seen from Table 6.7.1.1, there was a significant difference between KM and BCIF and between BCIG and BCIF, but not between KM and BCIG.

Table 6.7.1.1 - Summary of post-hoc tests: Time on since overshoot

	KM	**BCIG**
BCIG	t(17) = -0.302	
BCIF	t(17) = -7.716*	t(17) = -7.415*

*** = p < 0.05**

Inspection of the mean times indicated that KM (5.4 seconds) and BCIG (5.7 seconds) had significantly faster times than BCIF (11.4 seconds). Therefore, it could be concluded that after the target has been acquired, elimination when using KM and BCIG is significantly faster than eliminating a set of stationary targets while having to use facial expressions.

Through observation it was noticed that participants first completed one action before attempting the next when making use of BCIF. Thus, the cursor was first moved over the target, the head was then kept still and only thereafter did the participants attempt to form facial expressions in order to activate the shoot command. This may indicate that the multiple actions required from the same body part was the reason for the slower times. This, together with the fact that the participants were used to interacting with computers by only using their hands, caused increased difficulty for the participants.

Since the time on target metric gives a good indication of the time taken to activate a command, it can clearly be seen from the difference in mean times between BCIG and BCIF that the reason for the difference can be attributed to the method used for command activation. Therefore, it can be reasoned that the Peregrine glove was responsible for the faster activation of the shoot command. Consequently, KM and BCIG are better suited to activating the shoot command when dealing with a stationary target within a 2D game and can be recommended for use in this regard. When activating a command there is no significant difference between using a keyboard and the Peregrine glove, but when using facial expressions there is a significant difference, therefore indicating that the activation time when using facial expressions is much slower than the other two interfaces.

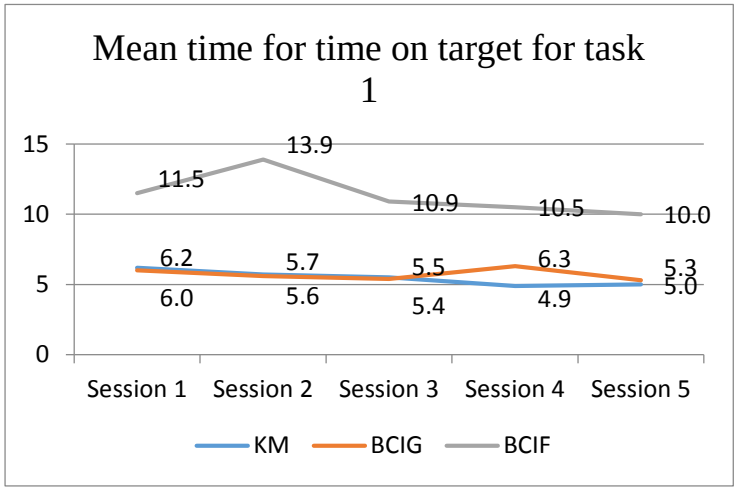

Chart 6.7.1.1 - Mean times for task 1

Participants' unfamiliarity with the concept of moving the cursor and activating a command through moving their head and face may be responsible for the slow command activation as seen with BCIF. As previously mentioned, the fact that the same part of the body is used, in this case the head using facial expressions and head movements, may explain the difficulty that participants faced with BCIF.

Sessions

At an α-value of 0.05 the null hypothesis $H_{0,2}$ can be rejected ($F(4, 204) = 3.086$, $p < 0.05$), indicating that there is a significant difference in the time the cursor spent on the target since the last overshoot from one session to the next.

Table 6.7.1.2 - Post-hoc tests: Time on target since overshoot

	Session 1	Session 2	Session 3	Session 4
Session 2	t(17) = 0.868			
Session 3	t(17) = 1.965	t(17) = 0.793		
Session 4	t(17) = 1.693	t(17) = 0.665	t(17) = -0.259	
Session 5	t(17) = 3.318*	t(17) = 2.347	t(17) = 1.45	t(17) = 1.939

*** = p < 0.05**

As can clearly be seen from Table 6.7.1.2, there was only a significant difference between sessions 1 and 5. There was a small decline in the mean time on target from each session to the next with the exception of session 1 to 2 (Chart 6.7.1.1). This indicates that learning took place. Part of the improvement was due to the participant becoming more familiar with the game, since improvement was present with KM as well. Therefore, it could be established that repeated use of the interfaces yielded improvements for the time on target metric.

When inspecting Chart 6.7.1.1, it is clear that the mean time for BCIF decreased from the earlier sessions to the later sessions. This result will be discussed in tasks 2 and 3 below.

6.7.2. Task 2: Time on target

In this instance, the assumption of sphericity ($\chi^2(9) = 21.167$, $p < 0.05$) was not met at an α-level of 0.05. A within-subjects repeated measures ANOVA indicated that there was significant interaction between the two factors, namely session and interfaces ($F(6.725, 171.500) = 5.256$, $p < 0.05$). Therefore, separate analyses had to be conducted by isolating each factor in turn.

Interfaces

Friedman tests showed a statistically significant difference between the different interfaces for all sessions except sessions 1 and 3 – Session 1: $\chi^2(2) = 2.333$, $p > 0.05$, Session 2: $\chi^2(2) = 19.000$, $p < 0.01$, Session 3: $\chi^2(2) = 0.333$, $p > 0.05$, Session 4: $\chi^2(2) = 16.444$, $p < 0.01$, Session 5: $\chi^2(2) = 32.444$, $p < 0.01$. Thus, the null hypothesis $H_{0,1}$ can be rejected for sessions 2, 4 and 5. Therefore, the interface used has an effect on the time on target metric for three of the five

sessions. Table 6.7.2.1 summarises the Wilcoxon signed rank test, post-hoc tests for the Friedman test, for each session and between each pair of interfaces:

Table 6.7.2.1 - Wilcoxon post-hoc: Time on target

	KM - BCIG	KM - BCIF	BCIG - BCIF
Session 2	Z = -1.198, p > 0.017	Z = -3.463, p < 0.017*	Z = -3.462, p < 0.017*
Session 4	Z = -0.719, p > 0.017	Z = -3.398, p < 0.017*	Z = -3.506, p < 0.017*
Session 5	Z = -3.005, p < 0.017*	Z = -3.724, p < 0.017*	Z = -3.724, p < 0.017*

*** = p < 0.017**

Inspection of the mean times for all 5 sessions indicated that KM (3.5 seconds) and BCIG (3.3 seconds) had significantly faster times than BCIF (4.7 seconds). Therefore, it could be concluded that eliminating predictable moving targets while making use of KM and BCIG is significantly faster than using BCIF. Consequently, KM and BCIG are better suited for this type of task and can be recommended for use in this regard.

Sessions

The time on target data for task 2 was found to be spherical for KM and BCIG but not for BCIF (Table 6.7.2.2).

Table 6.7.2.2 – Results of Mauchley's sphericity test and ANOVA

	KM	BCIG	BCIF
Sphericity test	χ^2 (9) = 12.256	χ^2 (9) = 4.454	χ^2 (9) = 19.112*
ANOVA	F (4, 68) = 1.891	F (4, 68) = 1.170	F (2.740, 46.582) = 7.297*

*** p < 0.05**

As can clearly be seen from Table 6.7.2.2, $H_{0,2}$ could be rejected for BCIF, consequently the sessions for BCIF had an effect on the time on target. Post-hoc tests were conducted to determine which sessions differed significantly from the other. The results of the post-hoc tests for BCIF are summarised in Table 6.7.2.3:

Table 6.7.2.3 - Post-hoc test results: Time on target after overshoot for BCIF

	Session 1	Session 2	Session 3	Session 4
Session 2	t(17) = 2.531			
Session 3	t(17) = 4.099*	t(17) = 2.094		
Session 4	t(17) = 3.957*	t(17) = 1.099	t(17) = -0.274	
Session 5	t(17) = 4.345*	t(17) = 1.61	t(17) = 0.39	t(17) = 1.056

*** p < 0.05**

As seen in Table 6.7.2.3 and Chart 6.7.2.1, there was a significant difference from the first session to the later sessions for BCIF. It is interesting to note that the time on target for BCIF notably increased from 3.6 during session 1 to 5.3 during session 5, as seen in Chart 6.7.2.1.

This is a different result to that of task 1, indicating that the moving target had an impact on the time on target metric. This may be explained by the observation that participants experienced difficulty while simultaneously attempting to move the cursor and activate the shoot command with BCIF. Due to participants ambushing the target and not actively tracking it, the contact times between the cursor and the target during earlier sessions were brief. The increased time spent on target could indicate that the participants grew more accustomed to

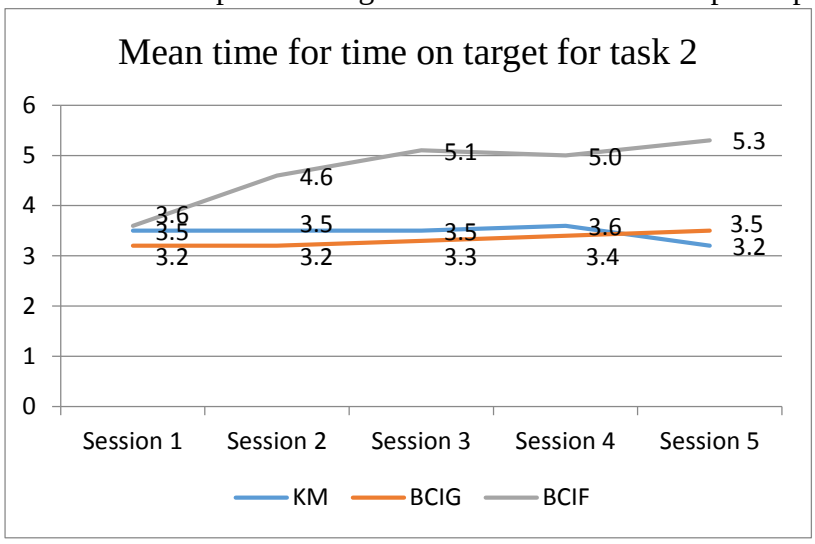

Chart 6.7.2.1 – Mean times for task 2

moving the cursor with their heads and were able to track the moving target while activating the shoot command by using their facial expressions. Thus, performing the moving and shooting actions simultaneously lead to longer times on target as a result of more stable cursor control. This indicates that learning did take place for one of the three interfaces, namely BCIF.

6.7.3. Task 3: Time on target

In this case, the assumption of sphericity (χ^2 (9) = 10.793, $p > 0.05$) was met at an α-level of 0.05. A within-subjects repeated measures ANOVA indicated that there was significant interaction between the two factors (F (8,204) = 9.878, $p < 0.05$), therefore separate analyses had to be conducted, thus separate ANOVAs and Friedman tests were conducted.

Interfaces

Friedman tests showed a statistically significant difference between the different interfaces for all sessions except session 1 and 2 – Session 1: $\chi^2(2) = 6.169$, $p < 0.05$, Session 2: $\chi^2(2) = 5.600$, $p > 0.05$, Session 3: $\chi^2(2) = 16.930$, $p < 0.01$, Session 4: $\chi^2(2) = 19.444$, $p < 0.01$, Session 5: $\chi^2(2) = 16.444$, $p < 0.01$. Therefore, the null hypothesis $H_{0,1}$ can be rejected for sessions 1, 3, 4 and 5. Thus, the interface used has an effect on the time on target metric for four of the five sessions. Table 6.7.3.1 summarises the Wilcoxon signed rank test, post-hoc tests for the Friedman test, for each session and between each pair of interfaces:

Table 6.7.3.1 - Wilcoxon post-hoc: Time on target

	KM – BCIG	**KM - BCIF**	**BCIG - BCIF**
Session 1	Z = -2.983, p < 0.017*	Z = -2.580, p < 0.017*	Z = -0.414, p > 0.017
Session 3	Z = -3.006, p < 0.017*	Z = -2.221, p > 0.017	Z = -3.375, p < 0.017*
Session 4	Z = -1.852, p > 0.017	Z = -3.201, p < 0.017*	Z = -3.680, p < 0.017*
Session 5	Z = -0.828, p > 0.017	Z = -3.506, p < 0.017*	Z = -3.506, p < 0.017*

*** = p < 0.017**

Inspection of the mean times over 5 sessions for task 3 indicated that KM (3.6 seconds) and BCIG (3.2 seconds) had significantly faster mean times than BCIF (4.3 seconds). Therefore, it could be concluded that targeting and eliminating unpredictable moving targets, while making use of KM and BCIG, are significantly faster than using BCIF. It is interesting to note that BCIG had a slightly faster mean time than the traditional KM, indicating that BCIG showed promise in terms of fast command activation.

Sessions

The data was found to be spherical for BCIG and BCIF but not for KM (Table 6.7.3.2). The results of the repeated measures ANOVA are tabulated in Table 6.7.3.2:

Table 6.7.3.2 - Results of Mauchley's sphericity test and ANOVA

	KM	**BCIG**	**BCIF**
Sphericity test	χ^2 (9) = 18.404*	χ^2 (9) = 3.726	χ^2 (9) = 10.978
ANOVA	F(2.436, 41.419) = 2.253	F (4, 68) = 2.222	F (4, 68) = 16.307*

*** p < 0.05**

As can clearly be seen from Table 6.7.3.2, $H_{0,2}$ could be rejected for BCIF but not for KM and BCIG, therefore the session for BCIF has an effect on the time on target metric. Since there were significant differences, post-hoc tests were conducted to determine which sessions

differed significantly from the other. The results of the post-hoc tests for BCIF are summarised in Table 6.7.3.3:

Table 6.7.3.3 - Post-hoc test results: Time on target after overshoot for BCIF

	Session 1	Session 2	Session 3	Session 4
Session 2	t(17) = 3.194			
Session 3	t(17) = 5.087*	t(17) = 2.631		
Session 4	t(17) = 7.511*	t(17) = 4.41*	t(17) = 0.442	
Session 5	t(17) = 6.069*	t(17) = 4.612*	t(17) = 0.651	t(17) = 0.452

*** p < 0.05**

As seen in Table 6.7.3.3 and Chart 6.7.3.1, there was a significant difference from the first session to the later sessions for BCIF, but not for KM and BCIG. Once again it is interesting to note that the time on target for BCIF notably increased from 3.1 during session 1 to 4.9 during session 5. As argued during task 2, this may not be an undesirable result as it may be reasoned that the participants grew more comfortable with moving the cursor with their heads and were able to actively track the moving targets while activating the shoot command by using the required facial expressions. Thus, it could be reasoned that learning took place during task 3 for BCIF.

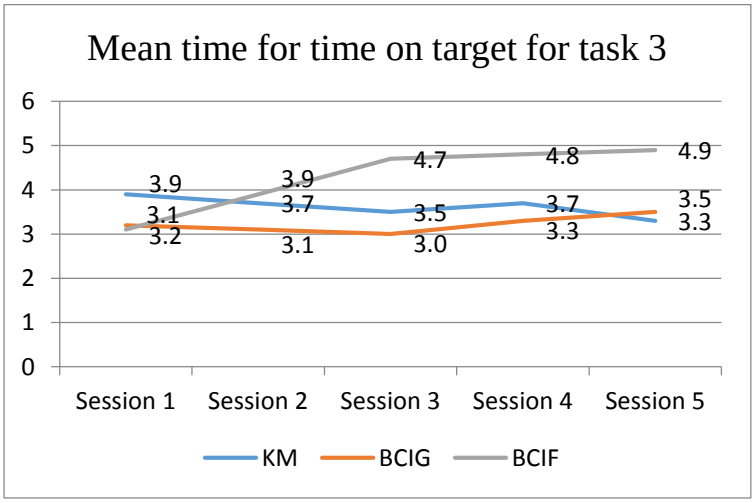

Chart 6.7.3.1 - Mean times for task 3

It would be interesting to investigate whether the increase in the time on target metric would have continued to increase with further sessions or whether it would have stabilised and then decreased as participants grew even more accustomed to using the interface and were able to activate the shoot command faster.

This indicates that BCIF was not very intuitive to use at first, but with some training the participants improved and were able to make effective use of the interface. This again leads to the idea of effective use of this interface by disabled users.

6.7.4. Time on target summary

The best interface to use for the shortest time on target before eliminating stationary targets was KM, with BCIG being the best for moving targets, as seen during tasks 2 and 3. This is an interesting result as the participants were not used to BCIG, thus indicating that BCIG was very intuitive and compared very well to the traditional KM, particularly in terms of the tasks that most closely resembled real gameplay. This indicated that participants could actively track and follow moving targets with the head-mounted mouse while simultaneously activating the shoot command using the glove. Further studies in terms of fast command activation can be undertaken in future by making use of the Peregrine glove in conjunction with another method of cursor movement.

Through observations it was noticed that participants performed better with the head-mounted mouse during tasks where the target was moving opposed to it being stationary. Participants experienced problems executing small head movements. Once they positioned the cursor close to the stationary target, they struggled to perform the last small movement to place it on the target. It was revealed that participants would rather move the cursor farther away from the stationary target and then reattempt to place it over the target. Since eye movement is used to track objects in everyday life (smaller movements) and head movement is only used to look in a different direction (larger movements) this is understandable. This also substantiates the result of the pilot study which showed that larger head movements (low sensitivity) produced the best results. As a result, it can be argued that it is difficult for participants to perform small head movements. This result is similar to that of Beelders (2011), where the individuals would rather move the cursor further away from the target, and then re-attempt the targeting action, when the cursor was very close to the target. It was reasoned that by using a gravitational well the cursor could be effectively pulled onto the nearest target once the cursor was within a certain distance from the target. Consequently, there will be no need for small adjustments of the cursor. Although Beelders (2011) utilised eye gaze as a method of input by, it can be reasoned that the recommendations made can also be applied to head-mounted cursor movement.

BCIF proved to be the slowest of the three interfaces, but it could still be successfully used as an interface when playing 2D games after several sessions. Interesting to note was the increase in time on target for BCIF from the earlier sessions to the latter. When comparing these results to the total time as well as the number of overshoots, an interesting result becomes apparent. The total time it took to complete the tasks as well as the number of overshoots from session 1 to 5 declined for all tasks, whereas the time on target increased. Thus, it can be reasoned that more stable active targeting resulted in longer times on target, but reduced the number of overshoots and thus also the total time to complete the tasks.

It was concluded that significant learning took place during all three tasks for BCIF, thus it can be reasoned that this interface was not very intuitive to use. Even so, this interface, with some training, could be successfully used by disabled users (individuals with little to no mobility below the neck) for 2D games.

6.8. Number of misses

The number of misses were measured by totalling the number of times the participant activated the shoot command without eliminating a target.

The following hypotheses were formulated for the analysis of the three tasks:

- $H_{0,1}$: The interface used has no effect on the number of misses.
- $H_{0,2}$: There is no difference in the number of misses between multiple sessions.

6.8.1. Task 1 and 2: Number of misses

In this instance, the assumption of sphericity was not met for either task 1 (χ^2 (9) = 30.136, $p < 0.05$) or task 2 (χ^2 (9) = 38.035, $p < 0.05$) at an α-level of 0.05. A repeated measures within-subjects ANOVA was performed to analyse the afore-mentioned hypotheses. In the case of significant difference in the time taken to complete the task, post-hoc tests were applied to determine which interfaces or sessions were responsible for the significant difference. Due to no significant interaction between the two factors for these two tasks, tasks 1 and 2 will be discussed under the same heading.

Interfaces

At an α-value of 0.05 the null hypothesis $H_{0,1}$ can be rejected for both task 1 ($F(2, 51) = 35.088$, $p < 0.05$) and task 2 ($F(2, 51) = 36.924$, $p < 0.05$), therefore the interface used has an effect on the number of misses. This result indicates that there is a significant difference in the number of times the participants activated the shoot command while not eliminating the target, when using the three user interfaces. The results for task 1 are shown on the first line of each cell and on the second line for task 2.

Table 6.8.1.1 - Summary of post-hoc tests: Number of misses

	KM	**BCIG**
BCIG	t(17) = 6.325, p < 0.05* t(17) = 2.889, p < 0.05*	
BCIF	t(17) = 7.919, p < 0.05* t(17) = 8.453, p < 0.05*	t(17) = 1.593, p > 0.05 t(17) = 5.565, p < 0.05*

*** = p < 0.05**

As can clearly be seen from Table 6.8.1.1, there was a significant difference between KM and BCIG as well as KM and BCIF, but not between BCIG and BCIF for task 1. During task 2 there was a significant difference between all three interfaces. Inspection of the mean number of misses for task 1 indicated that KM (1.1) had significantly fewer misses than BCIG (5.7) and BCIF (6.8). The same tendency was seen for task 2 where KM (3.2) had significantly fewer misses than BCIG (6.4) and BCIF (12.5).

There is a difference in the number of misses between tasks 1 and 2 for BCIF, but not for BCIG, indicating that the command activation was once again the cause for this difference. It could thus be argued that participants experienced more difficulty eliminating moving targets than stationary targets. This may also be a contributing factor to the large number of overshoots that were seen during tasks 2 and 3 for BCIF. As was previously discussed, through observation it was noticed that participants struggled to keep the cursor stable while attempting to activate the shoot command using facial expressions. The blinking required for this action resulted in small head movements, which caused the cursor to move off-target, where-after the shoot command was activated and therefore resulting in a miss. This was not the case when the Peregrine glove was utilised. Nevertheless, it could be concluded that eliminating the target when using KM is significantly more effective than eliminating a set of stationary targets or predictable moving targets while making use of BCIG or BCIF. Therefore, KM is better suited to effectively eliminate stationary and moving targets within a 2D game and can be recommended for use in this regard.

Sessions

At an α-value of 0.05 the null hypothesis $H_{0,2}$ can be rejected for task 1 (F (2.998, 5.996) = 9.486, $p < 0.05$) and task 2 (F (3.013, 153.663) = 16.157, $p < 0.05$). This result indicates that there is a significant difference in the number of misses for tasks 1 and 2 from one session to the next. Post-hoc tests were applied to determine which interfaces were responsible for the significant difference in the number of misses. The results for task 1 are shown on the first line and on the second line of each cell for task 2.

Table 6.8.1.2 - Summary of post-hoc tests: Number of misses sessions.

	Session 1	**Session 2**	**Session 3**	**Session 4**
Session 2	t(17) = -2.976* t(17) = -1.51		t(17) = -1.755 t(17) = 1.045	t(17) = 1.518 t(17) = 3.261*
Session 3	t(17) = -1.444 t(17) = -2.529	t(17) = 1.755 t(17) = -1.045		t(17) = 3.177* t(17) = 3.732*
Session 4	t(17) = -3.393* t(17) = -5.129*	t(17) = -1.518 t(17) = -3.261*	t(17) = -3.177 t(17) = -3.732*	
Session 5	t(17) = -4.382* t(17) = -6.391*	t(17) = -2.849 t(17) = -5.212*	t(17) = -5.415* t(17) = -5.894*	t(17) = -0.944 t(17) = -2.223

*** = p < 0.05**

As can clearly be seen from Table 6.8.1.2, there was a significant difference between some sessions for both tasks 1 and 2. There was a decline for all three interfaces in the time it took to complete task 1 from sessions 1 to 5, with the exception of session 3, where there was a slight increase.

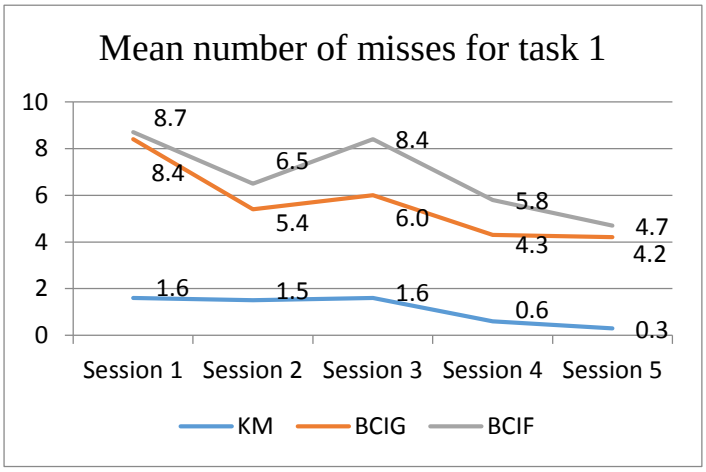

Chart 6.8.1.1 - Mean number of misses for task 1

There was a steady decline in the number of misses for task 2 from one session to the next, as can be seen in Chart 6.8.1.2. This indicates that learning took place. However, part of the improvement was due to the participant becoming more familiar with the game since improvement was present when using KM as well.

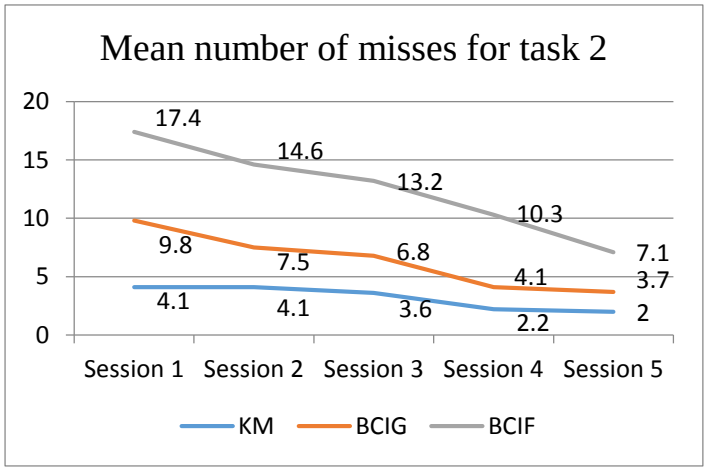

Chart 6.8.1.2 - Mean number of misses for task 2

Therefore, it could be established that repeated use of the interfaces yielded improvements for the number of misses for all interfaces.

6.8.2. Task 3: Number of misses

In this instance, the assumption of sphericity (χ^2 (9) = 46.646, $p < 0.05$) was not met at an α-level of 0.05. A repeated measures within-subjects ANOVA was performed to analyse the aforementioned hypotheses, which indicated significant interaction between the two factors, namely session and the interfaces ($F(5.241, 133.651) = 5.491$, $p < 0.05$).

Interfaces

Friedman tests showed a statistically significant difference between the different interfaces for all sessions – Session 1: $\chi^2(2) = 22.200$, $p < 0.01$, Session 2: $\chi^2(2) = 30.179$, $p < 0.01$, Session 3: $\chi^2(2) = 17.294$, $p < 0.01$, Session 4: $\chi^2(2) = 18.353$, $p < 0.01$, Session 5: $\chi^2(2) = 15.200$, $p < 0.01$. Therefore, the null hypothesis $H_{0,1}$ can be rejected for all five sessions. Table 6.8.2.1 summarises the Wilcoxon signed rank test, post-hoc tests for the Friedman test, for each session and between each pair of interfaces:

Table 6.8.2.1 - Summary of Wilcoxon post-hoc tests: Number of misses

	KM – BCIG	KM - BCIF	BCIG - BCIF
Session 1	Z = -3.482, p < 0.017*	Z = -3.682, p < 0.017*	Z = -1.564, p > 0.017
Session 2	Z = -3.116, p < 0.017*	Z = -3.733, p < 0.017*	Z = -3.359, p < 0.017*
Session 3	Z = -2.342, p > 0.017	Z = -3.662, p < 0.017*	Z = -3.212, p < 0.017*
Session 4	Z = -1.586, p > 0.017	Z = -3.465, p < 0.017*	Z = -3.140, p < 0.017*
Session 5	Z = -1.814, p > 0.017	Z = -3.656, p < 0.017*	Z = -1.985, p > 0.017

*** = p < 0.017**

As can clearly be seen from Table 6.8.2.1, there was a significant difference between all three interfaces for most sessions in terms of the mean number of misses during task 3.

Inspection of the mean number of misses indicated that KM (3.7) had significantly less misses than BCIG (7.7) as well as BCIF (13.0). The reason for the higher number of misses for BCIG and BCIF may be attributed to the time it took to activate the shoot command. The longer it takes to eliminate the target, the more time it has to move out of contact with the cursor, thus resulting in a miss. Therefore, it could be concluded that moving the cursor to make contact with the target and then activating the shoot command, while making use of KM, will result in significantly fewer misses than when using BCIG or BCIF. Consequently, KM is better suited to targeting and eliminating an unpredictable moving target. Furthermore, a significant difference in the mean times between BCIG and BCIF was also noted, indicating that BCIG would be more suitable than BCIF when faced with unpredictable moving targets.

Sessions

The data was found to be spherical for BCIF, but not for KM and BCIG (Table 6.8.2.2). The results of the repeated measures ANOVA are tabulated below:

Table 6.8.2.2 - Summary of Mauchley's sphericity test and ANOVA

	KM	BCIG	BCIF
Sphericity test	χ^2 (9) = 19.380*	χ^2 (9) =42.736*	χ^2 (9) = 7.106
ANOVA	F (2.378, 40.421) = 3.848*	F (1.681, 28.569) = 16.806*	F (4, 68) = 22.290*

*** p < 0.05**

As can clearly be seen from the table, $H_{0,2}$ could be rejected for all three interfaces, therefore the session has an effect on the number of misses for task 3. Post-hoc tests were conducted to determine which sessions differed significantly from the other. The results are summarised

below (KM is on the topmost row of each cell, BCIG directly below and BCIF on the last row of each cell):

Table 6.8.2.3 - Summary of post-hoc tests: Number of misses

	Session 1	Session 2	Session 3	Session 4
Session 2	t(17) = -1.492 t(17) = -3.997* t(17) = -2.374		t(17) = 0.676 t(17) = 1.197 t(17) = 2.446	t(17) = 1.706 t(17) = 3.237* t(17) = 3.57*
Session 3	t(17) = -1.849 t(17) = -3.637* t(17) = -5.246*	t(17) = -0.676 t(17) = -1.197 t(17) = -2.446		t(17) = 1.129 t(17) = 2.893 t(17) = 1.276
Session 4	t(17) = -2.234 t(17) = -4.864 t(17) = -5.399*	t(17) = -1.706 t(17) = -3.237 t(17) = -3.57*	t(17) = -1.129 t(17) = -2.893 t(17) = -1.276	
Session 5	t(17) = -3.869* t(17) = -5.429* t(17) = -7.443*	t(17) = -1.959 t(17) = -3.001 t(17) = -5.972*	t(17) = -1.715 t(17) = -2.338 t(17) = -5.636*	t(17) = -0.263 t(17) = 0.708 t(17) = -2.614

*** p < 0.05**

There was a steady decline in the number of misses for task 3 for all interfaces, from sessions 1 to 5, as seen in Chart 6.8.2.1. Inspection of the mean times indicated that the least number of misses for KM and BCIF was during session 5, but that the least number of misses took place during session 4 for BCIG.

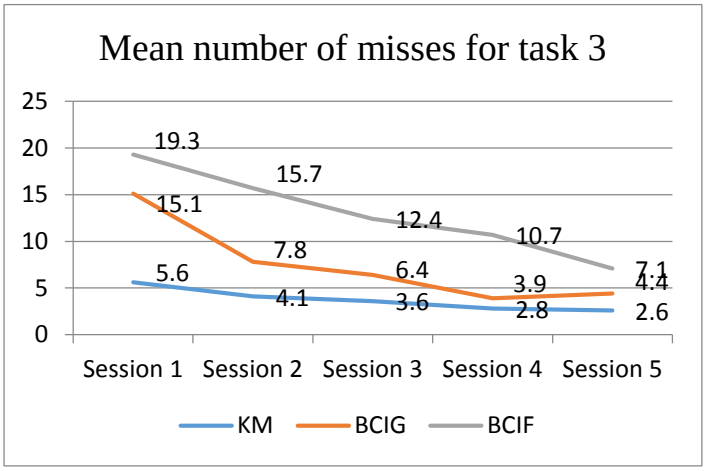

Chart 6.8.2.1 - Mean number of misses for task 3

From Chart 6.8.2.1, it can be seen that there was significant improvement from the first session to the later sessions for all three interfaces. This indicates that learning took place for the three interfaces, although the best improvement was noticed for BCIF. When inspecting the mean number of misses per session it can be seen that BCIF started with 19.3 overshoots during session 1, but improved with training to 7.1 misses during session 5. This indicates that extensive learning took place for BCIF, and it would be interesting to investigate whether the decline in the number of misses would have continued with further sessions. This indicates that

BCIF was not very intuitive to use, but with some training the participants were able to make effective use of the interface.

This again leads to the idea of effective use of this interface by disabled users (with limited or no use of arms and hands). Although not being very intuitive, with training it could pose to be a very effective interface.

6.8.3. Number of misses summary

The results of the analysis showed that KM was the best interface in terms of the number of misses for task 1, 2 and 3. This can be expected as this is the interface that the participants are very familiar with. During task 1, there was no significant difference between the two natural interfaces namely BCIG and BCIF, which was not seen during tasks 2 and 3. This indicates that BCIG and BCIF compared well when used to eliminate stationary targets. On the other hand, this is not the case for moving targets as seen during tasks 2 and 3. Thus, while both interfaces used the same method for cursor movement there is a significant difference in the mean number of misses when aiming at moving targets. This difference may be due to the time it takes to activate the shoot command with each interface as the target will move away from the cursor if the shoot command is not activated immediately, resulting in a miss. Therefore, a stationary target will result in fewer misses if the participants find it difficult to quickly activate the shoot command, as is the case with BCIF, where participants were observed struggling to perform both move and shoot actions simultaneously.

It was also noticed that when participants attempted to activate the shoot command by winking, they inadvertently moved their heads which led to movement of the cursor. This then led to the cursor moving off-target and resulting in a miss. This had an impact on both stationary and moving targets, but would be more pronounced for moving targets since the cursor and the target move simultaneously. When using the glove, participants could perform command activation and cursor movement simultaneously, which may explain the difference in the mean number of misses between the mentioned interfaces.

6.9. Non-eliminated targets

The number of non-eliminated targets was measured by totalling the number of targets that left the game window without being eliminated by the participant. This metric is only applicable

to tasks 2 and 3, as only moving targets can leave the game window. The high presence of zeros, in particular during the later sessions, prohibited statistical analysis of these metrics.

For KM, non-eliminated targets were mainly zero (there were two sessions that had a single, non-eliminated target for task 3). Therefore, participants could successfully eliminate all targets using the traditional interface. For BCIG task 2 there was a single, non-eliminated target during session 2, and 4 in session 1 of task 3, the remainder of the metrics were zero. The non-eliminated targets for BCIF are shown below:

Table 6.9.1 – Number of non-eliminated targets

	Session 1	**Session 2**	**Session 3**	**Session 4**	**Session 5**
BCIF – Task 2	56	11	6	0	0
BCIF – Task 3	58	27	15	6	0

BCIF improved over the sessions in terms of eliminating targets, with both task 2 and task 3 ending with zero non-eliminated targets. This closely resembles the metrics of number of overshoots and misses where the number of errors was initially very high, but then decreased as the participants grew more accustomed to the interface. It was found that participants often positioned the cursor over the target and then lost the target when attempting to eliminate it. In this instance, it shows that often they were unable to reposition the cursor before the target was out of bounds.

6.10. Resets

As discussed in Chapter 5, the reset function was introduced in order to rectify the misalignment between the position of the users' head and the cursor while making use of the head-mounted mouse.

The number of resets was measured by totalling the number of times the participants activated the reset command in order to re-align the cursor with their head position. This metric is only applicable to BCIG and BCIF as both interfaces make use of the head-mounted accelerometer for cursor control.

The following hypotheses were formulated for the analysis of the three tasks:

- $H_{0,1}$: The interface used has no effect on the number of resets.
- $H_{0,2}$: There is no difference in the number of resets between multiple sessions.

6.10.1. Task 1, 2 and 3

In this instance, the assumption of sphericity was not met for task 1 (χ^2 (9) = 91.351, $p < 0.05$), task 2 (χ^2 (9) = 104.172, $p < 0.05$) and task 3 (χ^2 (9) = 78.343, $p < 0.05$) at an α-level of 0.05, therefore the greenhouse-geisser adjustment was applied.

Interfaces

At an α-value of 0.05 the null hypothesis $H_{0,1}$ cannot be rejected for task 1 (F (1, 34) = 2.244, $p > 0.05$), task 2 (F (1, 34) = 1.606, $p > 0.05$) and task 3 (F (1, 34) = 2.567, $p > 0.05$), therefore, the interface used does not have an effect on the number of resets in any of the tasks.

Although there is no significant difference between the number of resets for BCIG and BCIF, inspection of the means indicated that BCIG (task 1 - 0.74, task 2 - 0.86, task 3 - 0.77) required less resets than BCIF (task 1 - 1.27, task 2 - 1.46, task 3 - 1.47). The difference, although not significant, indicates that an element of the interfaces was responsible for this difference. Both interfaces made use of the head-mounted accelerometer, thus the method for command activation may be the segregating factor. This further indicates that the use of facial expressions lead to more resets, especially with moving targets, which can be explained by the difficulties experienced by participants with the use of facial expressions.

As previously discussed, although the participants struggled with cursor control when making use of facial expressions, this was not the case when using the Peregrine glove. The necessity for activating the reset command is a direct result of poor cursor control when the participant moves the cursor off-screen and a misalignment between the participant's head and cursor is created. In order to correct the alignment, the reset command has to be activated.

Consequently, although there is no significant difference, BCIG had a slight advantage in terms of resets for all three tasks, indicating that BCIG would be the best choice to make in order to try and eliminate or moderate the need for the reset command.

Sessions

At an α-value of 0.05 the null hypothesis $H_{0,2}$ can be rejected for task 1 (F (2.108, 71.672) = 4.280, $p < 0.05$), task 2 (F (2.202, 74.868) = 6.381, $p < 0.05$) and task 3 (F (1.901, 64.621) = 6.598, $p < 0.05$), indicating that there is a significant difference in the number of resets from one session to the next for all three tasks. The results of the post-hoc tests for tasks 1, 2 and 3

are shown in Table 6.10.1, task 1 is shown on the first line of each cell, task 2 on the second, and task 3 on the third line.

Table 6.10.1 - Post-hoc tests: Number of resets

	Session 1	Session 2	Session 3	Session 4
Session 2	t(17) = 0.035 t(17) = -1.309 t(17) = -2.266			
Session 3	t(17) = -2.685 t(17) = -3.722* t(17) = -2.340	t(17) = -2.070 t(17) = -1.543 t(17) = -0.443		
Session 4	t(17) = -2.250 t(17) = -3.481* t(17) = -2.861	t(17) = -1.816 t(17) = -1.318 t(17) = -1.652	t(17) = 0.383 t(17) = 0.606 t(17) = -1.164	
Session 5	t(17) = -3.393* t(17) = -4.429* t(17) = -3.315*	t(17) = -2.687 t(17) = -2.166 t(17) = -3.194*	t(17) = -1.965 t(17) = -2.139 t(17) = -2.24	t(17) = -1.803 t(17) = -1.683 t(17) = -2.000

*** p < 0.05**

As can clearly be seen from Table 6.10.2.1, there was only a significant difference between sessions 1 and 5 for task 1. For task 2, there was a significant difference between session 1 and sessions 3, 4 and 5, and for task 3, significant differences between sessions 1 and 2 and session 5. The difference between task 1 (stationary targets) and tasks 2 and 3 (moving targets) may be explained by the fact that moving targets frequently moved close to the edge of the game window. It was more likely for the cursor to make contact with the edge of the game window if the participant was experiencing difficulties in eliminating the target. When the participant could not eliminate the target in time, the target moved off-screen, which the participant attempted to follow. This in turn caused a misalignment between the head and cursor position, which led to another reset. As can be seen from Chart 6.10.2.1, there was a decline in the mean time for BCIG with the exception of session 2 for task 1.

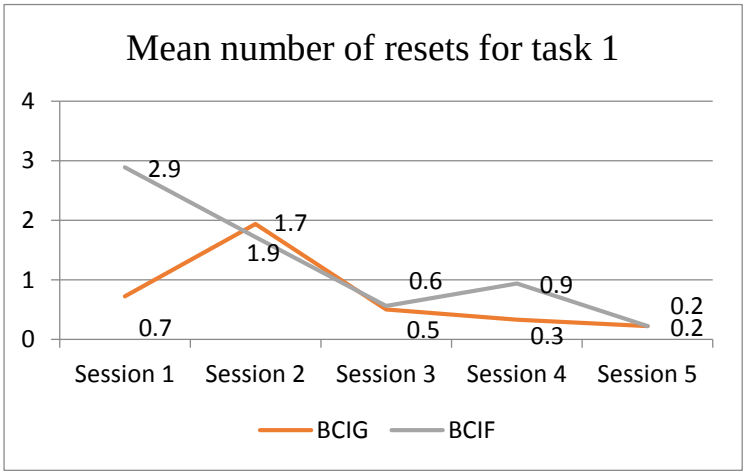

Chart 6.10.2.1 - Mean number of resets for task 1

This was also seen in the number of resets from each session to the next with the exception of session 4 for BCIF.There was a small decline in the mean number of resets from each session to the next with the exception of session 4 for task 2, as can be seen in Chart 6.10.2.2. The decline in the number of resets required was also seen in task 3, with a steady decline in the mean number of resets from each session to the next for both BCIG and BCIF, as can clearly be seen in Chart 6.10.1.3.

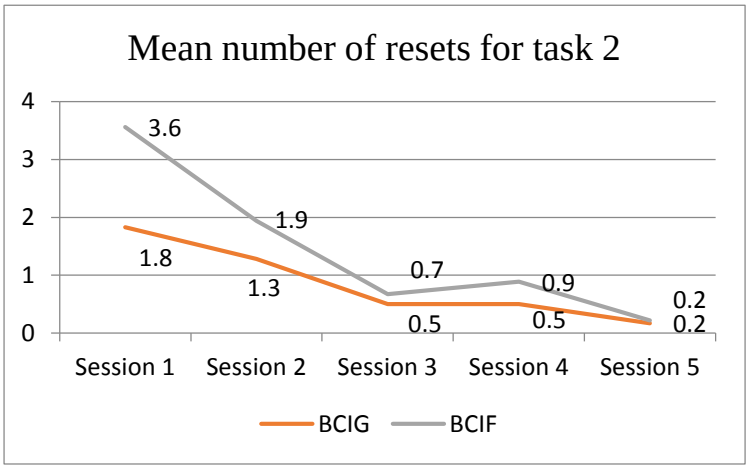

Chart 6.10.2.2 - Mean number of resets for task 2

The improvements that can be seen in the analysis of the number of misses and number of overshoots metrics can also be seen in the number of resets metric. Thus, there appears to be a correlation between the number of resets required and the participants' ability to control the cursor.

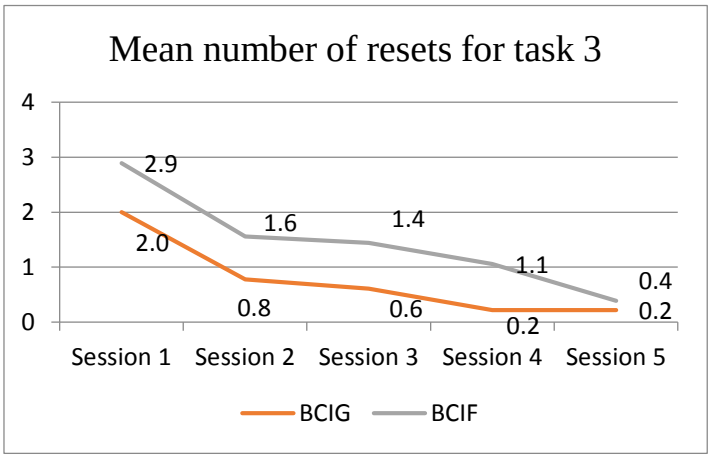

Chart 6.10.2.3 - Mean number of resets for task 3

As the participants' ability to control the cursor increased, the need for resets decreased, indicating that learning took place. Therefore, it could be established that repeated use of the interfaces yielded improvements for the number of resets metric. It is important to note that the

reset command is still required as the need for the reset command could not completely be eliminated during the 5 sessions.

6.11. Analysis of Subjective measures

Together with objective user testing, subjective measurements were also recorded by making use of questionnaires. These questionnaires included the QUIS (Appendix A) for BCIG and BCIF interfaces as well as the Pointing device assessment (Appendix B).

6.11.1. Satisfaction

The QUIS (Appendix A) questionnaire was used for both new interface combinations and was administered during the first and last session. The section of the QUIS that was used has 5 questions measured on a rating scale from 1 to 9.

Analysis of responses for BCIG and BCIF

The means and standard deviation of the responses to the five questions on the questionnaire are shown in Table 6.11.1.1. The results of the questionnaire analysis show that these responses are statistically significant overall for BCIG (Z = -3.553, p = 0.001) as well as BCIF (Z = -3.334, p = 0.001).

Table 6.11.1.1 - Means and standard deviation of session 1 and 5 for BCIG and BCIF.

			BCIG		BCIF	
	1	**9**	**Session 1**	**Session 5**	**Session 1**	**Session 5**
1	Terrible	Wonderful	$\bar{x}$ = 6.7 (sd = 1.3)	$\bar{x}$ = 8.6 (sd = 0.6)	$\bar{x}$ = 4.7 (sd = 2.4)	$\bar{x}$ = 7.7 (sd = 1.4)
2	Difficult	Easy	$\bar{x}$ = 6.6 (sd = 1.7)	$\bar{x}$ = 8.7 (sd = 0.5)	$\bar{x}$ = 3.5 (sd = 1.8)	$\bar{x}$ = 7.4 (sd = 1.4)
3	Frustrating	Satisfying	$\bar{x}$ = 6.1 (sd = 2.1)	$\bar{x}$ = 8.4 (sd = 1.9)	$\bar{x}$ = 4.0 (sd = 2.1)	$\bar{x}$ = 7.4 (sd = 2.0)
4	Dull	Stimulating	$\bar{x}$ = 7.4 (sd = 1.6)	$\bar{x}$ = 8.6 (sd = 0.7)	$\bar{x}$ = 6.0 (sd = 2.4)	$\bar{x}$ = 7.9 (sd = 1.2)
5	Rigid	Flexible	$\bar{x}$ = 7.0 (sd = 1.5)	$\bar{x}$ = 8.4 (sd = 0.7)	$\bar{x}$ = 4.6 (sd = 2.4)	$\bar{x}$ = 7.6 (sd = 1.4)

Individual question analysis showed that all five questions had significant differences in responses between the two sessions (Table 6.11.1.2) for both interfaces.

Table 6.11.1.2 – Results of Wilcoxon tests

Question	Session 1 - Session 5 for BCIG	Session 1 - Session 5 for BCIF
1	Z = -3.558, p = 0.000*	Z = -2.906, p = 0.004*
2	Z = -3.447, p = 0.001*	Z = -3.624, p = 0.001*
3	Z = -2.634, p = 0.008*	Z = -3.012, p = 0.003*
4	Z = -2.243, p = 0.025*	Z = -2.418, p = 0.016*
5	Z = -2.823, p = 0.005*	Z = -2.969, p = 0.003*

For BCIG, during session 5, the participants rated question 1 higher (8.6) than during the first session (6.7). Although the mean score during session 1 showed that the participants initially enjoyed BCIG, they enjoyed the interface more during the last session, as the mean was very close to the maximum score of nine that can be assigned to the question. This tendency was present in the other four questions as well. Thus, after five sessions, the participants felt that the interface was easier, more satisfying, more stimulating and more flexible to use. Consequently, it can be reasoned that participants initially enjoyed using BCIG and after several sessions their level of satisfaction further increased. Therefore, the responses during the first session indicates that BCIG was intuitive to use, further strengthening the results seen in the analysis of the usability data, which also showed that the interface was natural and intuitive to use.

In contrast, when inspecting the means for question 1, 3 and 5 for BCIF, all three questions had initial responses ranging between 4 and 5. This changed during the last session where all the responses were between 7 and 8. This indicated that the participants experienced a slightly negative feeling towards BCIF initially, which over several sessions improved to a positive overall response. Therefore, participants felt that BCIF was initially more terrible than wonderful, also more frustrating than satisfying, and more rigid than flexible to use. After repeated use of the interface, these negative feelings changed to positive responses from the participants.

It is interesting to note that although both BCIG and BCIF make use of the same device for the purpose of pointing there is a noticeable difference in the responses from the participants. This can be explained by the different methods used for command activation, namely the Peregrine glove for BCIG and facial expression for BCIF. By inspecting the responses for both interfaces it becomes clear that participants were much more at ease with BCIG than with BCIF, thus it can be argued that the use of facial expressions was responsible for the initial lack of satisfaction with BCIF. It could also be reasoned that not only the facial expression on its own, but also a combination of the head movement required for pointing combined with facial expressions for command activation, may be the cause of the initial difficulty experienced. This further strengthens the argument made during the analysis of objective measures where it was reasoned that the close proximity of the two modalities led to the difficulty experienced during the initial sessions.

When inspecting the mean response for question 4, it is quite encouraging to see an initial mean response of 6.0, which then improved to 7.9 during the last session. Although participants felt BCIF was difficult and frustrating to use during the initial sessions, they thought that the interface was interesting and stimulating to use. This can be regarded as a favourable factor for BCIF since the interface requires initial training before it becomes usable. The extra time required for training may cause the users to abandon the interface. Thus, it can be reasoned that because of the initial positive response to question 4, participants might spend the extra time on training because they find the interface interesting and stimulating.

During the analysis of objective measures it was argued that this interface could be used by disabled users (with little to no control of arms and hands) after some initial training. The fact that the participants found the interface stimulating when they initially used it, further encourages the idea that BCIF could be used as an alternative method of input for disabled users.

Pointing Device Assessment Questionnaire (Appendix B)

In this case the pointing device used is the head-mounted accelerometer attached to the Emotiv BCI, which was used for both NUIs. The questionnaire was administered during the first and last session in order to estimate the satisfaction level of the participants and if it changed as their exposure to the interface increased. The same method of data analysis used by Douglas, Kirkpatrick and MacKenzie (1999) was used during the analysis of the questionnaire data.

The means and standard deviation of the responses on the nine questions on the questionnaire are shown in Table 6.11.1.3. The results of the questionnaire analysis show that these responses are statistically significant overall ($Z = -3.041$, $p = 0.002$). It is important to note that comparisons in the same row between session 1 and 5 is meaningful but not between different rows since the directionality of the questions in the questionnaire differed.

Table 6.11.1.3 - Means and standard deviation for device assessment questionnaire

Question	Session 1	Session 5
1	$\bar{x}$ = 2.8 (sd = 0.7)	$\bar{x}$ = 2.5 (sd = 0.8)
2	$\bar{x}$ = 3.6 (sd = 0.8)	$\bar{x}$ = 4.2 (sd = 0.9)
3	$\bar{x}$ = 2.9 (sd = 1.1)	$\bar{x}$ = 2.5 (sd = 1.0)
4	$\bar{x}$ = 2.7 (sd = 1.1)	$\bar{x}$ = 2.0 (sd = 1.0)
5	$\bar{x}$ = 2.9 (sd = 1.3)	$\bar{x}$ = 1.6 (sd = 0.6)
6	$\bar{x}$ = 3.1 (sd = 0.6)	$\bar{x}$ = 2.8 (sd = 0.4)
7	$\bar{x}$ = 2.5 (sd = 1.1)	$\bar{x}$ = 1.7 (sd = 1.0)
8	$\bar{x}$ = 3.4 (sd = 1.0)	$\bar{x}$ = 3.9 (sd = 1.2)
9	$\bar{x}$ = 3.9 (sd = 1.0)	$\bar{x}$ = 4.5 (sd = 0.9)

Individual question analysis showed that question 2 - pertaining to the smoothness, question 4 - the physical effort, question 5 - referring to accurate pointing, question 7 - neck fatigue and question 9 - on the overall ease of use, had significant differences in the responses between the first and last sessions (Table 6.11.1.4).

Table 6.11.1.4 - Results of Wilcoxon tests

Question	Analysis results
1	Z = -1.897, p = 0.58
2	Z = -2.368, p = 0.018*
3	Z = -1.502, p = 0.133
4	Z = -2.106, p = 0.035*
5	Z = -3.194, p = 0.001*
6	Z = -1.311, p = 0.190
7	Z = -2.122, p = 0.034*
8	Z = -1.588, p = 0.112
9	Z = -2.396, p = 0.017*

* $p < 0.05$

When inspecting the means, as seen in Table 6.11.1.3, it can be extrapolated that there was a positive change in the responses of participants. This has the following implications: the participants felt that the smoothness during the operation of the pointing device was better during session 5. Additionally, the participants felt that they needed to exert less effort to utilise the pointing device and that accurate pointing was easier compared to session 1. An important finding was that participants noticed less neck fatigue as they grew more accustomed to the idea of moving the cursor by making use of head movements. Since this study, in part, investigated interfaces that will reduce RSI, it is a welcome finding that participants indicated that there was reduced neck fatigue after some use of the pointing device. The mean response for session 1 was 2.5, which indicates that participants were unsure about the amount of neck fatigue as the mean response is neutral on the rating scale of 5. During session 5, the

participants had a clearer view on the subject, which indicated that they experienced little neck fatigue after some use of the device. Furthermore, participants indicated that the overall ease of use of the pointing device improved through repeated use.

These results imply that participants grew more accustomed to, and accepting, of the interface as time went by. Thus, it can be reasoned that repeated use of the interface resulted in the increase in the level of user satisfaction. This improvement can be partly explained by the participants' initial lack of experience with the interface. During the first session this may have caused participants to feel uncomfortable with the use of the interface. After several sessions, as their effective use of the interface increased, the participants' experience of the interface showed a similar trend.

This result confirms the objective metric analysis of first contact time metric and time on target metric. Both these metrics mainly focused on the use of the pointing device during the specific tasks. It was found that learning took place, consequently more effective and efficient use of the pointing device was possible after several sessions.

6.12. Conclusion

This chapter discussed the analysis of the individual metrics for this study, as well as the responses from the post-test questionnaires. KM was found to be the best interface to use for all metrics, except during tasks 2 and 3 for the time on target metric, where BCIG produced the best results. Excluding the time on target metric, BCIG was found to be the second best interface with BCIF being the least promising interface to use. BCIG was the most intuitive interface, as the users' hands were used to perform actions using the glove, which they are more accustomed to. This was also confirmed in the responses to the post-test questionnaires. The responses indicated that the participants enjoyed BCIG more than BCIF, both initially and during later sessions. Although BCIF initially performed very poorly in most of the metrics, participants were able to make successful use of the interface after several sessions. The responses from the post-test questionnaire indicated that after some training, participants enjoyed BCIF much more than during the first session.

The next chapter will provide a summary of the results of this study and how these results will be used to answer the research questions posed. It will also include a discussion on the recommendations that can be made.

Chapter 7 - Conclusion

7.1. Introduction

The previous chapter discussed the results of the statistical analysis of the user testing. Chapter 7 will conclude the research by providing a summary of the research results and addressing each of the research questions initially posed. This chapter will include a discussion of the usability results in terms of effectiveness, efficiency, learnability and satisfaction. The recommendations of the study as well as suggestions for further research will also be discussed.

7.2. Motivation

In the domain of console gaming, alternative input technologies are available and have been accepted by console gamers. These technologies include input devices such as the Sony PlayStation and Nintendo Wii, where users make use of a motion controller to issue commands. Conversely, while most PC based gamers still utilise the traditional keyboard and mouse combination when playing games (Beckhaus, Blom & Haringer, 2005), this input method has stagnated over the past years (Reimer, 2005). Not only has the traditional user interface not evolved, but it leads to RSI, which is caused by frequent repetitions of monotonous actions (Kumar, 2014). Therefore, the need arises for alternative input technologies that will eliminate the need for learned actions, where the latter should rather be replaced by intuitive actions for control during gameplay. The next step in the evolution of PC gaming input may therefore lie in the field of NUIs (Reimer, 2005). The NUIs that were investigated during the study included the Emotiv BCI and the Peregrine gaming glove. The Emotiv's head-mounted accelerometer was used for cursor control in both NUIs.

Furthermore, benefits or difficulties that could be experienced by gamers while making use of these NUIs may be revealed. This may lead to some light being shed on why there has been little change in the way gamers interact with their computers.

7.3. Aim

The aim of this study was to investigate, within a gaming environment, the usability of two multimodal NUIs. These included the use of facial expressions as well as glove input for command activation, which were combined with a head-mounted accelerometer for cursor

control. These two NUIs were compared to a traditional, bimanual interface, namely the keyboard and mouse.

7.4. Results

In order to answer the research question posed in Chapter 1: To what extent is the usability of a game influenced through the use of a NUI as opposed to a traditional keyboard and mouse combination, the four hypotheses (from Chapter 1) that relate to usability have to be accepted or rejected. Each hypothesis will be addressed in the applicable subsection:

- $H_{0,1}$: The interface used has no effect on the effectiveness of task completion in a 2D game.
- $H_{0,2}$: The interface used has no effect on the efficiency of task completion in a 2 two-dimensional (2D) game.
- $H_{0,3}$: There is no difference between the learnability of a NUI and a GUI.
- $H_{0,4}$: The interface used has no effect on the level of satisfaction of the user in a 2 two-dimensional (2D) game.

Several other questions were posed during this study, the most important being: why have PC gamers not migrated to more natural intuitive interfaces for gameplay? In an attempt to answer this question another question was posed: is there a difference between the usability of NUIs and the traditional keyboard and mouse combination? If so, may this be the reason why widespread use of NUIs is not prevalent in PC gaming?

User testing was conducted in order to address these questions. The metrics that were used during the testing can be divided into different categories according to the usability model proposed by Nielsen (2012), namely effectiveness, efficiency, learnability and satisfaction.

Effectiveness

In order to measure the effectiveness of each interface during this study the non-eliminated targets metric (number of targets that left the game window without being eliminated) was used. This metric indicated that KM was the most effective interface, closely followed by BCIG.

BCIG compared well with KM in terms of effectiveness, although it could not better the traditional interface combination. Although initially being the least effective interface of the

three, BCIF showed the same level of effectiveness as KM and BCIG in session 5, indicating that it is initially less intuitive to use. Although not being very natural and intuitive to use this interface opens up new possibilities for a certain group of users. Due to both cursor movement, as well as command activation, being controlled by movement of the head and face it creates the opportunity for use by disabled users, especially individuals who do not have adequate control over their hands and arms. Consequently, with some training disabled users could successfully utilise this interface for 2D games and general point-and-click tasks. BCIF can also be used by this group for day-to-day computer tasks, which involve moving the cursor over items such as icons and then activating a command in order to run a certain application. These actions are similar to those found in task 1 (where the participant was required to eliminate a set of stationary targets), where, after some initial difficulties, participants were able to make effective use of BCIF.

The two NUIs that have been investigated could not better the traditional KM combination. It could be reasoned that gamers who are already accustomed to KM would not invest in a new interface which will at best match the effectiveness of KM. In the area of gaming, every small advantage that a gamer can get is important, thus using an interface that does not operate effectively is not a viable option.

Therefore, the hypothesis $H_{0,1}$ can be rejected as it is clear from the above results that there is a difference in the effectiveness of the three interfaces, with KM proving to be the most effective.

Efficiency

The efficiency of the different interfaces were compared by making use of the time measurements as well as the number of errors. The time measurements include the total time to complete the task, reload time, initial contact time as well as time on target metrics, whereas the number of errors will be represented by the number of overshoots and number of misses metric. When taking all these metrics into consideration KM seems to be the most efficient interface to use for gaming, while BCIG was the second most efficient and BCIF the least efficient.

This result could be expected as most individuals are accustomed to using KM for PC input. Although BCIG and BCIF made use of the same action for cursor movement, a difference in

terms of efficiency can be seen. Following is a discussion on a few of the observations made that could possibly explain this difference.

Participants were unfamiliar with the concept of simultaneously moving the cursor and activating a command by using facial expressions and head movement, which essentially uses the same body part. This may indicate that the multiple actions required from the same body part was the reason for the lower levels of efficiency for BCIF.

This dual action that had to be performed by using the same body part, led to the problem of participants unable to simultaneously perform cursor movement as well as command activation. Either cursor movement or command activation was first completed before the other was attempted. With BCIG, command activation and cursor movement could be performed simultaneously by the participants from the first session. It may be reasoned that participants performed better with BCIG due to the familiarity of using finger actions, since the participants were proficient with the keyboard and mouse, which primarily makes use of finger actions.

Additionally, it was observed that when participants attempted to perform the shoot action while using BCIF, it resulted in small movements of the head. This unintended head movement caused the cursor to move off target and resulted in overshoots as well as misses. This observation confirms that of Logsdon (2011), where the same part of the body (the users' hand) was used for command activation as well as cursor control. In Section 2.2.5 it was reasoned that this problem could be eliminated by making use of a different body part to control the cursor. During observation it was noticed that the difficulty experienced with the facial expressions in conjunction with head movement was not present when the command activation was conducted using the Peregrine glove. Thus, by making use of a different body part, the small unintended movements when attempting to activate a command were eliminated.

BCIG had the lowest time on target for moving targets, while KM had the lowest for stationary targets. This metric measured how fast a command could be activated. This is an exciting result as the participants were not used to BCIG, therefore indicating that BCIG was rather intuitive to use and compared very well to the traditional KM interface with fast command activation. Since this metric only measures command activation, it can be reasoned that the Peregrine glove is responsible for these positive results. Consequently, the Peregrine glove in combination with another method for cursor control may prove to be a promising research option for fast command activation in games.

Through observations it was noticed that participants, while using the head-mounted mouse, performed better during tasks where the target was moving, opposed to it being stationary. Additionally, it was observed that participants would rather move the cursor farther away from the stationary target when small adjustments are required and then reattempt to position the cursor over the target. Thus, participants experienced difficulty when executing small head movements, confirming the results by Beelders (2011) where participants experienced difficulty when attempting to perform small adjustments to the cursor position. Although Beelders (2011) used eye gaze as input, it may be reasoned that the solution suggested in the study can be used for head-mounted mouse control as well. This solution will be further discussed in Section 7.5.

As discussed, there is a difference between the level of efficiency of the NUIs selected for this study and the traditional combination of the keyboard and mouse. Consequently, the hypothesis $H_{0,2}$ can be rejected as it is clear that there is a difference in terms of efficiency for the three interfaces. With KM being the most efficient, it can be argued that with efficiency being a very important factor in gameplay, individuals would not migrate to an interface that offers them less in terms of efficiency than what they currently have.

Therefore, it may be reasoned that more research and development is required in order for NUIs to be accepted as efficient alternative interfaces for gameplay.

Learnability

The improvement of performance in terms of all the metrics indicated that learnability is present with these interfaces. In order to investigate learnability participants were tested over 5 sessions. The analysis of the different metrics revealed that learning took place for all three interfaces, thus there was an element of learning the game and not just the interface.

The three interfaces did not show similar levels of improvement, with KM showing the least, followed by BCIG, while the best improvement was noticed for BCIF. It could be reasoned that KM, which the participants were used to, only improved as far as learning of the game was concerned. BCIG showed improvement, although not as much as BCIF, indicating that BCIG was more intuitive to use than BCIF. BCIG and BCIF used the same action for cursor movement, therefore, the difference lies in the action used for command activation or in a combination of the method of command activation as well as cursor control.

The fact that repeated use of BCIF was required before BCIG and BCIF could be used at a similar level of accuracy indicates that BCIF is not as intuitive as initially thought, whereas BCIG did not require extensive learning before it could be used with acceptable levels of accuracy. The analysis of the subjective measures revealed that, although participants thought that BCIF was difficult and frustrating to use during the initial session, they felt that the interface was interesting and stimulating to use. It could be argued that this will be in favour of BCIF, which requires initial training before it can be used successfully. Therefore, it can be reasoned that participants may spend the extra time required for training due to the initial response that the interface was interesting and stimulating to use. As a result, BCIG, although not being very intuitive, after some training becomes a viable input option.

Thus, $H_{0,3}$ can be rejected since there is a difference in the learnability of the three interfaces that were investigated.

Satisfaction

The validity and reliability of subjective measurements where ensured by using established, existing questionnaires to gather data. The QUIS questionnaires (Harper & Norman, 1993), in conjunction with the Pointing Device Assessment Questionnaire (ISO 9241, 2000), was used for this purpose.

By analysing the subjective measurements, it was found that participants initially enjoyed using BCIG, and after several sessions their level of satisfaction improved. Thus, the responses during the first session indicate that the participants thought that BCIG was intuitive to use.

In contrast, the participants initially had a slightly negative feeling towards BCIF, which then improved over several sessions to an overall positive response. This initial negative response to the interface may be explained by the method used for command activation, namely facial expressions. By inspecting the responses for both interfaces it becomes clear that participants were much more at ease with BCIG than with BCIF. Thus, the use of facial expressions may be responsible for the initial lack of satisfaction with BCIF. Therefore, it can be argued that the participants had initially experienced greater difficulty in using BCIF than BCIG. However, after 5 sessions, participants indicated that they had an overall positive experience with both BCIG and BCIF, although some training was first required for BCIF.

The responses to the pointing device (head-mounted accelerometer) assessment questionnaire showed improvements in the level of satisfaction with the interface over several sessions. The responses during the first session were neutral, indicating that participants were neither negative nor positive about the pointing device. Participants' responses became more approving after several sessions. It is interesting to note that both BCIG and BCIF make use of the same action for the purpose of pointing, however, there is a noticeable difference in the responses from the participants. It can thus be reasoned that the participants were of the opinion that they enjoyed using the pointing device, therefore, the problem may be the method of command activation. The cause for this lack of satisfaction may be found in the close proximity of the actions for both command activation and cursor movement in BCIF, as previously argued.

Therefore, $H_{0,4}$ can be rejected as it is clear that there was a difference in the opinions voiced by the participants in terms of the NUIs, the participants were satisfied with BCIG from session one as opposed to BCIF where the participants initially were not satisfied, which improved after several sessions.

Since all four hypotheses have been rejected, and it can clearly be seen from the results above that KM was the most effective and most efficient of the interfaces investigated, as well as the difference noted in learnability and level of satisfaction, the research question can now be answered. The usability of a game is influenced through the use of a NUI, specifically, the usability in this case is negatively affected by the use of the two NUIs when compared to the traditional keyboard and mouse combination. When comparing the interfaces of the two NUIs it was noticed that there was a difference between the usability of BCIG and BCIF, with BCIG proving to be the better of the two. Since the same method of cursor control was used in both NUIs, this indicates that the Peregrine glove was more usable than the facial expressions offered by the Emotiv BCI.

7.5. Recommendations

The KM combination has been the main input device for many years and has been shown to be more effective and efficient than the NUIs in this study. For individuals to accept and migrate to a more natural interface the new interface will have to provide more effective and efficient input than what is already achievable with KM. A large obstacle in the way of alternative input methods is the existing skill and acceptance that computer users have with KM.

The difference in effectiveness and efficiency between a GUI and NUI was the central focus of the study. Since it was found that the traditional keyboard and mouse combination was more effective and efficient than the NUIs investigated, it is thus recommended that KM be used instead of the two NUIs for the most effective and efficient 2D gaming input.

Nonetheless, the Peregrine gaming glove has been found a suitable method for fast command activation in games, and combining the glove with an alternate pointing device may provide more successful results than those found in this study. It could be reasoned that the positive results achieved by the Peregrine glove may be attributed to the fact that it was the only device that made use of the users' hands and fingers to input instructions. The BCIG, in particular the Peregrine glove, appears to be a natural method of input as it required little to no training before it could be successfully used.

The attempt to maintain the stability of the cursor while activating a command when using the same appendage was problematic, as it negatively affected the effectiveness and efficiency of BCIF. The close proximity of the command activation and cursor control for BCIF posed a problem, therefore, it is recommended that the triggers for these actions do not make use of the same modality.

Furthermore, BCIF could be used by disabled individuals, who do not have adequate control of their arms and hands, for daily computer tasks, since the interface only makes use of head and facial movement for command activation and cursor movement. However, initial training with BCIF is recommended before daily use is attempted as there is an element of learning required.

The participants also experienced difficulty moving the cursor small distances with the head-controlled mouse. A solution to this problem may be to make use of a gravitation well as argued by Beelders (2011), where the cursor is pulled onto the target and kept there when the cursor is placed within a certain distance from the target. It is also recommended that a reset action be made available when using the head-mounted accelerometer due to the misalignment that can occur when the user turns their head too far, as discussed in Chapter 6. Alternatively, a probable solution to this problem could be to unbind the cursor from the game window. This would allow the cursor to move freely outside the game window and still stay aligned with the user's head position, therefore eliminating the misalignment.

Although the NUIs in this study could not better the KM in terms of the metrics used, the participants expressed their interest in the NUIs, and thus it can be recommended that the NUIs in this study can be used to make the gaming experience more stimulating and exciting. This would, however, not be recommended for individuals who are serious gamers who require the fastest and most accurate input in order to be competitive in the game that they are playing.

7.6. Further research

After the results of this study have been analysed, several recommendations for future studies can be made:

- In future research, real gameplay can be used during user testing, instead of tasks that were created for the study.
- Facial expressions may be used in conjunction with another method of cursor control due to the difficulty experienced with the use of the same modality for both command activation and cursor control. This may lead to more promising results.
- As a result of the positive outcome in terms of fast command activation, further studies can be undertaken by making use of the Peregrine glove in conjunction with another method of cursor movement.
- Additionally, more sessions could be added to the user testing to investigate if the improvements noticed would continue over several more sessions.
- The majority of the participants in this study had not used other input devices except KM before participating in the study. It would be interesting to replicate the study but with individuals with no computer training. This would give a clearer comparison between the rates of learning for the three interfaces, as prior exposure to the traditional interface KM would not be a factor then.
- Due to memorability being excluded from this study future research may include this metric.
- This study focused on NUIs in a 2D gaming environment - 3D gaming may be the next environment to investigate.

7.7. Summary

This study investigated NUIs for gameplay. It was revealed that the traditional user interface combination was more effective and efficient than the NUIs that were tested. Although individuals indicated that they enjoyed using the NUIs after several sessions, the objective

measurements indicated that the NUIs did not compare well in terms of the metrics (effectiveness, efficiency, learnability and satisfaction) for the study.

The results of this study indicated that there is a difference between the usability of the traditional input combination, the keyboard and mouse, and the two NUIs investigated in this study. KM was responsible for far more effective and efficient input, with one exception being fast command activation when using the Peregrine glove, where the two interfaces compared well. It may be argued that the difficulty experienced with the widespread acceptance of NUIs for computer gameplay may be the fact that the traditional interface combination is well entrenched and established as a reliable input combination.

In order for NUIs to replace the traditional input combination, the NUIs will have to offer faster and more effective input than what can be achieved with the traditional input combination. Merely matching the input accuracy and speed of the keyboard and mouse may not be enough to sway the minds of computer gamers.

It may therefore be reasoned that the use of KM has become a stumbling block to the adoption of NUIs, since KM has become natural and intuitive to computer users than the actual NUIs. Conversely, it can be reasoned that KM may in fact be the first natural user interface, as it extends the primary method of interaction that individuals employ on a daily basis to manipulate their immediate environment, namely with their hands.

Taking into consideration the results of this study, an argument can be made that not all alternative interfaces are natural interfaces. Therefore, it is essential that more research be conducted to first identify the actual natural methods of interaction that individuals employ to influence their environments on a daily basis before new NUIs are designed and developed. These actual natural methods of interaction can then be used as the foundation for new devices or combination of devices that will not merely match but exceed that which is achievable with the traditional keyboard and mouse combination. In closing, it is concluded that the two NUIs, when used for gameplay, caused a decrease in usability compared to the traditional mouse and keyboard combination. This indicates that additional research has to be conducted to improve the usability of the NUIs that were investigated in order for them to be comparable to the traditional interface combination during gameplay.

References

Abdi, H. (2010). The greenhouse-geisser correction. *Encyclopedia of Research Design*, 1-10.

Abran, A., Khelifi, A., Suryn, W., & Seffah, A. (2003). Consolidating the ISO Usability Models. *Proceedings of 11th international software quality management conference*.

Adamovich, S. V., Merians, A. S., Boian, R., Tremaine, M., Burdea, G. S., Recce, M., & Poizner, H. (2004). A virtual reality based exercise system for hand rehabilitation post-stroke: transfer to function. *In the Proceedings of the 26th Annual International Conference of the IEEE In Engineering in Medicine and Biology Society. 2*, pp. 4936-4939. IEEE.

Adamo-Villani, N., & Wilbur, R. (2007). An immersive game for K-5 math and science education. *In the Proceedings of the 11th International Conference on Information Visualization* (pp. 921–924). IEEE.

Adnan, N. H., Wan, K., Shahriman, A. B., Zaaba, S. K., Razlan, Z. M., Hazry, D., Aziz, A. (2012). Gesture Recognition Based On Hand Postures And Trajectories By Using Dataglove: A Fuzzy Probability Approach–A Review. *In the Proceedings of the International Conference on Man-Machine Systems.*

Akilli, G. K. (2005). User satisfaction evaluation of an educational website. *The Turkish Online Journal of Educational, 4*(1), 85-92.

Alejandro, J., & Nicu, S. (2007). Multimodal human-computer interacttion: A Survey. *Computer Vision and Image Understanding, 108*(1), 116-134.

Allison, B. (2011). Trends in BCI research: progress today, backlash tomorrow. *The ACM Magazine for Students - Neuroscience and Computing: Technology on the Brain, 18*(1), 18-22.

Allison, B. Z., & Pineda, J. A. (2003). ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system. *Neural Systems and Rehabilitation Engineering, 2*(11), 110-113.

Anderer, P., Roberts, S., Schlögl, A., Gruber, G., Klösch, G., Herrmann, W., Saletu, B. (1999). Artifact Processing in Computerized Analysis of Sieep EEG - A Review. *Neuropsychobiology, 40*, 150-157 .

Baecker, R., Grudin, J., Buxton, W. A., & Greenberg, S. (2014). *Readings in Human-Computer Interaction: toward the year 2000.* Morgan Kaufmann.

Bailey, K. (2008). *Methods of social research.* Simon and Schuster.

Beckhaus, S., Blom, K. J., & Haringer, M. (2005). A New Gaming Device and Interaction Method for a First-Person-Shooter. Leipzig : Proceedings of the Computer Science and Magic.

Beelders, T. R. (2011). *Enhancing the User Experience for a Word Processor Application Through Vision and Voice.* Doctoral dissertation, University of the Free State.

Benjamin, J., & Keller, R. M. (2003). Enabling Computer Decisions Based on EEG Input. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 4*(11), 354-360.

Bérard, F. (1999). The perceptual window: Head motion as a new input stream. Grenoble: In the proceedings of the IFIP Conference on human-computer interaction.

Bevan, N., & Macleod, M. (1994). Usability measurement in context. *Behaviour and Information Technology, 13*, 132-145.

Biggs, G., & MacDonald, B. (2003). A survey of robotic programming systems. *In the Proceedings of the Australasian conference on robotics and automation*, (pp. 1-3).

Blankertz, B., Sannelli, C., Halder, S., Hammer, E., Kübler, A., Müller, K., Dickhaus, T. (2009). Predicting BCI Performance to Study BCI Illiteracy. *In the proceedings of BMC Neuroscience 2009.* Berlin, Germany.

Boot, L. (2009). *Facial expressions in EEG/EMG recordings.* University of Twente : Master's thesis.

Bryson, S., & Levit, C. (1992). The virtual wind tunnel. *IEEE Computer graphics and Applications, 12*(4), 25–34.

Burke, A., & Pepper, E. (2002). Cumulative trauma disorder risk for children using computer products: results of a pilot investigation with a student convenience sample. *Public health reports, 117*(4), 350.

Calder, B. J., Phillips, L. W., & Tybout, A. M. (1982). The Concept of External Validity. *Journal of Consumer Research, 9*(3), 240-244.

Camara, A. (2011). Natural User Interfaces. 13th IFIP TC 13 international conference on Human-computer interaction-Volume Part I: Springer-Verlag.

Camastra, F., & De Felice, D. (2013). LVQ-based hand gesture recognition using a data glove. In *Neural Nets and Surroundings* (pp. 159-168). Berlin Heidelberg: Springer.

Cato, J. (2001). *User-centered web design.* Great Britain: Addison-Wesley.

Cechanowicz, J., & Gutwin, C. (2009). Augmented Interactions: A Framework for Adding Expressive Power to GUI Widgets. Uppsala: Human-Computer Interaction – INTERACT 2009.

Cernea, D., Olech, P., Ebert, A., & Kerren, A. (2011). EEG-based measurement of subjective parameters in evaluations. *In the Proceedings of HCI International 2011–Posters' Extended Abstracts* (pp. 279-283). Berlin Heidelberg: Springer.

Constantine, L. L., & Lockwood, L. A. (1999). *Software for Use: A Practical Guide to the Models and Methods of Usage-Centred Design.* New York: Addison-Wesley.

Coyle, D., Principe, J., Lotte, F., & Nijholt, A. (2013). Guest Editorial: Brain/Neuronal–Computer Game Interfaces and Interaction. *IEEE Transactions on Computational Intelligence and AI in Games, 5*(2), 77-81.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage publications.

Curran, E. A., & Stokes, M. J. (2003). Learning to control brain activity: A review of the production and control of EEG components for driving brain–computer interface (BCI) systems. *Brain and Cognition, 51*(3), 326-336.

da Silva, G. A., Nogueira, P. A., & Rodrigues, R. (2014). Multimodal vs. unimodal biofeedback in videogames: An empirical player study using a First-Person Shooter. Barcelona: Information Systems and Technologies (CISTI), 2014 9th Iberian Conference IEEE.

Deller, M., Ebert, A., Bender, M., & Hagen, H. (2006). Flexible gesture recognition for immersive virtual environments. *In the Proceedings of the Tenth International Conference on Information Visualization* (pp. 563–568). IEEE.

Diftler, M. A., Culbert, C. J., Platt, R., & Bluethmann, W. J. (2003). Evolution of the NASA/DARPA robonaut control system. *In the Proceedings of the IEEE International Conference on Robotics and Automation. 2*, pp. 2543–2548. IEEE.

Dipietro, L., Sabatini, A. M., & Dario, P. (2008). A survey of glove-based systems and their applications. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 38*(4), 461-482.

Doerr, K. U., Rademacher, H., Huesgen, S., & Kubbat, W. (2007). Evaluation of a low-cost 3D sound system for immersive virtual reality training systems. *IEEE Transactions on Visualization and Computer Graphics, 13*(2), 204–212.

Dollman, G. J. (2014). *An Exploration of the use of Brain-Computer Interfaces for Human-Robotic Interaction: A comparative study using the Emotiv headset against traditional control techniques.* Master's thesis, University of the Free State.

Donovan, J., & Brereton, M. (2005). Movements in gesture interfaces. *In the Workshop Call for Participantion*, (p. 6).

Douglas, S. A., Kirkpatrick, A. E., & MacKenzie, I. S. (1999). Testing pointing device performance and user assessment with the IS0 9241, Part 9 Standard. *In Proceedings of CHI '99*, (pp. 215-222). Pennsylvania, United States of America.

Dowdy, S., & Weardon, S. (1983). *Statistics for research.* New York: John Wiley and Sons.

Dumas, J. S., & Redish, J. C. (1999). *A practical guide to usability testing* (Revised Edition ed.). Portland: Intellect books.

Durlach, N., & Mavor, A. (1995). *Virtual Reality: Scientific and Technological Challenges.* Washington, DC: National Academies Press.

Emotiv Inc. (2015). *Emotiv EPOC*. Retrieved 01 29, 2015, from Emotiv: http://www.emotiv.com/epoc/

Evans, P. T., Vance, J. M., & Dark, V. J. (1999). Assessing the effectiveness of traditional and virtual reality interfaces in spherical mechanism design. *Journal of Mechanical Design, 121*(4), 507–514.

Fatourechi, M., Bashashati, A., Ward, R. K., & Birch, G. E. (2007). EMG and EOG artifacts in brain computer interface systems: A Survey. *Clinical Neurophysiology, 118*, 480–494.

Faulkner, C. (1998). *The essence of human-computer interaction.* Great Britain: Prentice-Hall.

Field, A., & Hole, J. G. (2002). *How to design and report experiments.* Sage.

Finke, A., Lenhardt, A., & Ritter, H. (2009). The MindGame: a P300-based brain-computer interface game. *Neural Network, 22*, 1329–1333.

Foottit, J., Brown, D., Marks , S., & Connor , A. M. (2014). An Intuitive Tangible Game Controller. *In the Proceedings of the 2014 Conference on Interactive Entertainment.* New York.

Frankfort-Nachmias, C., & Nachmias, D. (1992). *Research Methods in the Social Sciences* (4th ed.). London: Edward Arnold.

Freeman, W. T., & Weissman, C. D. (1994). *Television Control by Hand Gestures.* Technical Report, Mitsubishi Electronic Research Laboratories, TR-94-24.

Galán, F., Ferrez, P., Oliva, F., Guardia, J., & del R Millan, J. (2007). Feature extraction for multi-class BCI using canonical variates analysis. *Proceedings of the IEEE International Symposium on Intelligent Signal Processing* (pp. 1-6). IEEE.

Ganin, I. P., Shishkin, S. L., & Kaplan, A. Y. (2013). A P300-based braincomputer interface with stimuli at moving objects: four-session singletrial and triple-trial tests with a game-like task design. *PloS one, 8*(10), e77755.

Gao, W., Ma, J., Shan, S., Chen, X., Zeng, W., Zhang, H., Wu, J. (2000). Handtalker: A multimodal dialog system using sign language and 3-d virtual human. *In the Proceedings of the International Conference on Advances Multimodal Interfaces* (pp. 564–571). Berlin Heidelberg: Springer.

Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., Edlinger, G. (2009). How many people are able to control a P300-based brain–computer interface (BCI). *Neuroscience Letters, 1*(462), 94-98.

Guger, C., Edlinger, G., Harkam, W., Niedermayer, I., & Pfurtscheller, G. (2003). How many people are able to operate an EEG-based brain-computer interface. *IEEE transactions on rehabilitation engineering, 11*(2), 145-147.

Gurkok, H., Hakvoort, G., & Poel, M. (2011). Evaluating User Experience in a Selection Based Brain-Computer Interface Game A Comparative Study. *In the Proceedings of the 10th international Conference on Entertainment Computing* (pp. 77-88). Springer.

Gurkok, H., Nijholt, A., & Poel, M. (2012). Brain-Computer Interface Games: Towards a Framework. *In the Proceedings of Entertainment Computing-ICEC 2012* (pp. 373 - 380). Berlin Heidelberg: Springer.

Harper, B. D., & Norman, K. L. (1993). Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5. *In the Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference*, (pp. 224-228).

Heger, D., Putze, F., & Schultz, T. (2011). Online recognition of facial actions for natural EEG-based BCI applications. *In Proceedings of Affective Computing and Intelligent Interaction. 6975*, pp. 436-446. Berlin Heidelberg: Springer.

Heise, D. R. (1970). The Semantic Differential and Attitude Research. In G. F. Summers (Ed.), *Attitude Measurement* (pp. 235-253). Chicago: Rand McNally.

Helten, T., Baak, A., Muller, M., Bharaj, G., Seidel, H. P., & Theobalt, C. (2013). Personalization and evaluation of a real-time depth-based full body tracker. *In proceedings of the 2013 International Conference on 3D Vision - 3DV* (pp. 279-286). IEEE.

Heo, H., Lee, E. C., Park, K. R., Kim, C. J., & Whang, M. (2010). A Realistic Game System Using Multi-Modal User Interfaces. *IEEE Transactions on Consumer Electronics, 56*(3), 1364-1372.

Hernandez-Rebollar, J. L., Kyriakopoulos, N., & Lindeman, R. W. (2002). The AcceleGlove: A whole-hand input device for virtual reality. *In the Proceedings of ACM SIGGRAPH conference on abstracts and applications* (p. 259). ACM.

Hewett, T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G., Verplank, W. (1996). *ACM SIGCHI Curricula for Human-Computer Interaction.* ACM SIGCHI.

Hinkelmann, K., & Kempthorne, O. (2007). *Design and Analysis of Experiments: Introduction to Experimental Design.* Wiley.

Hix, D., Templeman, J. N., & Jacob, R. J. (1995). Pre-Screen Projection: From Concept to Testing of a New InteractionTechnique. *In the Procedings of the the ACM CHI'95 Human Factors in Computing Systems Conference*, pp. 226-233.

Hondrou, C., & Caridakis, G. (2012). Affective, Natural Interaction Using EEG: Sensors, Application and Future Directions. In *Artificial Intelligence: Theories and Applications.* (pp. 331–338). Berlin Heidelberg: Springer-Verlag .

Hooper, C. J., & Dix, A. (2012). Web Science and Human-Computer Interaction: When Disciplines Collide. Proceedings of the 3rd Annual ACM Web Science Conference. ACM.

Hoste, L., Dumas, B., & Signer, B. (2012). SpeeG: a multimodal speech-and gesture-based text input solution. *In the Proceedings of the International working conference on advanced visual interfaces* (pp. 156-163). ACM.

Howard, B., & Howard, S. (2001). Lightglove: Wrist-worn virtual typing and pointing. *In the Proceedings of the Fifth International Symposium on Wearable Computers* (pp. 172–173). IEEE.

Hwang, H. J., Kwon, K., & Im, C. H. (2009). Neurofeedback-based motor imagery training for brain–computer interface (BCI). *Journal of neuroscience methods, 179*(1), 150-156.

Iron Will Innovations. (2013). *DotA2 and LOL Guide for Peregrine Glove.* Retrieved 03 17, 2015, from http://theperegrine.com/wp-content/themes/peregrine/downloads/LoL_Basic_Guide_1.pdf

Isaac, S., & Michael, W. (1997). *Handbook in research and evaluation: A collection of principles, methods and strategies useful in the planning, design amd evaluation of studies in education and the behavioural sciences* (3rd ed.). San Diego: Educational and Industrial Testing Services.

ISO 9241. (2000). *Ergonomic requirements for office work with visual display terminals (VDTs) - Requirements for non-keyboard input devices.*

ISO 9126. (2001). ISO Standard 9126: Software Engineering – Product Quality, parts 1, 2 and 3. In *International Organization for Standarization.* Geneve.

Jackson, M., & Mappus, R. (2010). Applications for Brain Computer Interfaces. In *Brain Computer Interfaces* (pp. 89-103). London: Springer.

Jacob, R. J. (1993). Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces. *Advances in Human Computer Interaction, 4*, 151-190.

Jacob, R. J. (1996). Input Devices and Techniques. In A. B. Tucker, & A. B. Tucker (Ed.), *The Computer Science and Engineering Handbook* (pp. 1494-1511). CRC Press.

Jacob, R. J., Leggett, J. J., Myers, B. A., & Pausch, R. (1993). Interaction styles and input/output devices. *Behaviour & Information Technology, 12*(2), 69-79.

Jain, J., Lund, A., & Wixon, D. (2011). The Future of Natural User Interfaces. *In CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 211-214). Vancouver: ACM.

Jimenez, L., Andujar, M., & Gilbert, J. (2012). *Reader Engagement.*

Johannsen, G. (2009). Human Machine Interaction. *Control systems, Robotics and Automation, 21*.

Kaikkonen, A., Kekäläinen, A., Canka, M., Kallio, T., & Kankainen, A. (2005). Usability testing of mobile applications: a comparison between laboratory and field testing. *Journal of Usability Studies, 1*(1), 4-16.

Karray, F., Alemzadeh, M., Saleh, J. A., & Arab, M. N. (2008). Human-Computer Interaction: Overview on State of the Art. *International Journal on Smart Sensing and Intelligent Systems, 1*(1), 137-159.

Keefe, D. F., Feliz, D. A., Moscovich, T., Laidlaw, D. H., & LaViola, J. J. (2001). Cave painting: A fully immersive 3D artistic medium and interactive experience. *In Proceedings of the 2001 symposium on Interactive 3D graphics* (pp. 85–93). ACM.

Kennedy, P. R., Bakay, R. A., Moore, M. M., Adams, K., & Goldwaithe, J. (2000). Direct Control of a Computer from the Human Central Nervous System. *IEEE Transactions on Rehabilitation Engineering, 8*(2), 198-202.

Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2001). The analysis of repeated measures designs: A review. *British Journal of Mathematical and Statistical Psychology, 54*(1), 1-20.

Kester, J. M. (2012). *Emotiv as an Alternative Input Device for the Physically Disabled.* Oswego: University of New York.

Key, J. P. (1997). *Research Design in Occupational Education.* Oklahoma: Oklahoma State University.

Kim, J., Bee, N., Wagner, J., & Andre, E. (2004). Emote to Win: Affective Interactions with a Computer Game Agent. *GI Jahrestagung, 1*, 159-164.

Kjeldskov, J., & Stage, J. (2004). New techniques for usability evaluation of mobile systems. *International journal of human-computer studies, 60*(5), 599-620.

Kramer, J. P., Lindener, P., & William , R. G. (1989). *Unites States of America Patent No. 5,047,952.*

Krapichler, C., Haubner, M., Loesch, A., & Englmeier, K. H. (1997). A human–machine interface for medical image analysis and visualization in virtual environments. *19In the Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. 4*, pp. 2613–2616. IEEE.

Kristensson, P. O., & Vertanen, K. (2011). Asynchronous Multimodal Text Entry Using Speech and Gesture Keyboards. *In the Proceedings of INTERSPEECH*, (pp. 581-584).

Kübler, A., Nijboer, F., Mellinger, J., Vaughan, T. M., Pawelzik, H., Schalk, G., Wolpaw, J. R. (2005). Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology, 10*(64), 1775-17777.

Kumar, P., Verma, J., & Prasad, S. (2012). Hand Data Glove: A Wearable Real-Time Device for Human - Computer Interaction. *International Journal of Advanced Science and Technology, 43*, 15-25.

Kumar, S. (2014). Musculoskeletal Discomforts in Body Posture of Computer Users: A Quantitative Analysis. *Industrial Engineering Letters, 4*(10), 31-39.

Kuroda, T., Tabata, Y., Goto, A., Ikuta, H., & Murakami, M. (2004). Consumer price data-glove for sign language recognition. *In the Proceedings of the International Conference for Disability : Virtual Reality Associated Technology*, (pp. 253–258). Oxford.

Lal, T. N., Hinterberger, T., Widman, G., Schroder, M., Hill, J., Rosenstiel, W., Birbaumer, N. (2005). Methods Towards Invasive Human Brain Computer Interfaces. In L. K. Saul, Y. Weiss, & L.

Bottou (Eds.), *Advances in Neural Information Processing Systems (NIPS)* (Vol. 17, pp. 737–744). Cambridge: MIT Press.

LaViola, J. J. (1999). *A survey of hand posture and gesture recognition.* Providence: Brown University Technical Report.

Lazar, J., Feng, J. H., & Hochheiser, H. (2010). *Research methods in human-computer interaction.* John Wiley & Sons.

Lebedev, M. A., & Nicolelis, M. A. (2006). Brain–machine interfaces: past, present and future. *Trends in Neuroscience, 29*(9), 536–546.

Lee, B., Isenberg, P., Riche, N. H., & Carpendale, S. (2012). Beyond mouse and keyboard: Expanding design considerations for information visualization interactions. *IEEE Transactions on Visualization and Computer Graphics, 18*(12), 2689-2698.

Leeb, R., Perdikis, S., Tonin, L., Biasiucci, A., Tavella, M., Creatura, M., Millán, J. R. (2013). Transferring brain–computer interfaces beyond the laboratory: successful application control for motor-disabled users. *Artificial intelligence in medicine, 59*(2), 121-132.

Levy, P. S., & Lemeshow, S. (2013). *Sampling of Populations: Methods and Applications.* John Wiley & Sons.

Lewis, J. R. (1995). Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *International Journal of Human-Computer Interaction, 7*(1), 57-78.

Lewis, J. R., Potosnak, K. M., & Magyar, R. L. (1997). Keys and Keyboards. In *Handbook of human computer interaction.* (pp. 1285–1315). Amsterdam.

Li, F. W., Lau, R. W., & Ng, F. F. (2003). VSculpt: A distributed virtual sculpting environment for collaborative design. *IEEE Transactions on Multimedia, 5*(4), 570–580.

Li, S. Y., Cox, A. L., Blandford, A., Cairns, P., & Abeles, A. (2006). Further investigations into post-completion errors: the effects of interruption position and duration. *Proceedings of the 28th Annual Meeting of the Cognitive Science Society*, (pp. 471-476).

Logsdon, D. A. (2011). *Arm-Hand-Finger Video Game Interaction.* Texas: A&M University.

Lotte, F. (2011). Brain-Computer Interfaces for 3D Games: Hype or Hope. *Proceedings of the 6th International Conference on Foundations of Digital Games* (pp. 325-327). ACM.

MacKenzie, I. S. (2012). *Human-computer interaction: An empirical research perspective.* Toronto: Newnes.

Malizia, A., & Belucci, A. (2012). The Artificiality of Natural User Interfaces. *Communications of the ACM*, pp. 36-38.

Manresa, C., Perales, F. J., Mas, R., & Varona, J. (2005). Hand tracking and gesture recognition for human-computer interaction. *In ELCVIA: electronic letters on computer vision and image analysis, 5*, 96-104.

Marchall, M. N. (1996). Sampling for qualitative research. *Family Practice, 13*, 522-525.

Marklin, R. W., Simoneau, G. G., & Monroe, J. F. (1999). Wrist and forearm posture from typing on split and vertically inclined computer keyboard. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 4*(41), 559-569.

McFarland, D. J., Krusienski, D. J., Sarnacki, W. A., & Wolpaw, J. R. (2008). Emulation of computer mouse control with a noninvasive brain-computer interface. *Journal of Neural Engineering, 5*(2), 101-110.

McGuffin, M., & Balakrishnan, R. (2005). Fitts' law and expanding targets: Experimental studies and design for user interfaces. *ACM Transactions on Computer-Human Interaction, 12*(4), 388-422.

Miao, Y., Metze, F., & Rawat, S. (2013). Deep maxout networks for low-resource speech recognition. *In the Proceedings of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding* (pp. 398-403). IEEE.

Minke, A. (1997). Conducting repeated measures analyses: Experimental design considerations. *In the Proceedings of the Annual Meeting of the Southwest Educational Research Association*, 23-25.

Minoh, M., Obara, H., Funtatomi, T., Toyura, M., & Kakusho, K. (2007). Direct manipulation of 3D virtual objects by actors for recording live video content. *In the Proceedings of the Second International Conference on Informatics Research for Development of Knowledge Society Infrastructure* (pp. 11–18). IEEE.

Molich, R., & Jeffries, R. (2003). Comparative expert reviews. *In the Proceedings of the Extended Abstracts on Human Factors in Computing Systems* (pp. 1060-1061). ACM.

Morita, H., Hashimoto, S., & Ohteru, S. (1991). A computer music system that follows a human conductor. *Computer, 7*(24), 44-53.

Mouton, J. (2001). *How to succeed in your Master's and Doctoral Studies* (1st ed.). Pretoria: Van Schaik.

Moyle, M., & Cockburn, A. (2005). A Flick in the Right Direction: A Case Study of Gestural Input. *Behaviour and Information Technology, 24*(4), 275-288.

Mugler, E. M., Ruf, C. A., Halder, S., Bensch, M., & Kubler, A. (2010). Design and implementation of a P300-based brain-computer interface for controlling an internet browser. *Neural Systems and Rehabilitation Engineering, 6*(18), 599-609.

Mulder, A., Fels, S., & Mase, K. (1997). Empty-handed gesture analysis in MAX/FTS. *In the Proceedings of the AIMI international workshop on Kansei-the technology of emotion*, (pp. 87–90).

Murakami, K., & Hitomi, T. (1991). Gesture Recognition Using Recurrent Neural Networks. *In the Proceedings of CHI'91 Human Factors in Computing Systems* (pp. 237-242). ACM.

Myeung-Sook, Y., Joonho, K., & Sunghoon, K. (2010). NeuroWander: a BCI Game in the Form of Interactive Fairy Tale. *In Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing-Adjunct* (pp. 389-390). ACM.

Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain Computer Interfaces, a Review. *Sensors, 12*, 1211-1279.

Niedermeyer, E., & da Silva, F. (2005). *Electroencephalography: basic principles, clinical applications, and related.* Lippincot Williams & Wilkins.

Nielsen, J. (1992). The usability engineering life cycle. *Computer, 25*(3), 12-22.

Nielson, J. (1993). Noncommand User Interfaces. *Communications of the ACM, 36*(4), 83-99.

Nielsen, J. (1994). *Usability engineering.* Elsevier.

Nielsen, J. (2000). *Why You Only Need to Test with 5 Users.* Retrieved 03 10, 2015, from http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

Nielsen, J. (2012). *Usability 101: Introduction to Usability*. Retrieved 08 12, 2013, from http://www.nngroup.com/articles/usability-101-introduction-to-usability/

Nielsen, J., & Landauer, T. (1993). A mathematical model of the finding of usability problems. *In the Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 206-213). Amsterdam: ACM.

Nijholt, A., Tan, D., Pfurtscheller, G., Brunner, C., Millán , J., Allison, B., Popescu, F. (2008). Brain-computer interfacing for intelligent systems. *Intelligent Systems, 3*(23), 76–83.

Nintendo. (2014). *Nintendo Virtual Console*. Retrieved 04 21, 2015, from http://www.nintendo.com/games/detail/yEiAzhU2eQI1KZ7wOHhngFoAHc1FpHwj

Oikonomidis, L., Kyriazis, N., & Argyros, A. (2011). Efficient Model-based 3D Tracking of Hand Articulations using kinect. *BMVC, 2*(1), 3.

Olivier, M. (2004). *Information technology research: A practical guide for Computer Science and Informatics* (2nd ed.). Pretoria: Van Schaik.

Orozco, M., El Saddik, A., Petriu, E., & Silva, J. (2012). *The role of haptics in games.* INTECH Open Access Publisher.

Oviatt, S., & Cohen, P. (2000). Perceptual user interfaces: multimodal interfaces that process what comes naturally. *Communications of the ACM, 43*(3), 45 - 53.

Oviatt, S., & VanGent, R. (1996). Error Resolution During Multimodal Human-Computer Interaction. *Proceedings of the International Conference on Spoken Language Processing* (pp. 204-207). IEEE.

Pao, L., & Speeter, T. H. (1989). Transformation of human hand positions for robotic hand control. *In the Proceedings of the IEEE International Conference on Robotics and Automation. 3*, pp. 1758–1763. IEEE.

Peregrine. (2013). *Peregrine*. Retrieved 01 26, 2015, from http://theperegrine.com/

Pereira, A., Laroche, C., Rempel, D., & Hsieh, C. (2014). The Effect of Keyboard Key Spacing on Typing Speed, Error, Usability, and Biomechanics Part 2. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 56*(4), 752-759.

Peshock, A., Dunne, L. E., & Duvall, J. (2014). Argot: A Wearable One-handed Keyboard Glove. *In the Proceedings of the 2014 ACM International Symposium on Wearable Computers: Adjunct Program* (pp. 87-92). Seattle, USA: ACM.

Piekarski, W., & Thomas, B. H. (2001). Tinmith-Metro: New outdoor techniques for creating city models with an augmented reality wearable computer. *In the Proceedings of the Fifth International Symposium on Wearable Computers* (pp. 31-38). IEEE.

Plass-Oude Bos, D., Reuderin, B., van de Laar, B., Gürkök, H., Mühl, C., Poel, M., Heylen, D. (2010). Brain-Computer Interfacing and Games. In D. S. Tan, & A. Nijholt (Eds.), *Brain-Computer Interfaces* (pp. 149-178). London: Springer-Verlag.

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. *Encyclopedia of human computer interaction, 1*, 211-219.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human Computer Interaction.* Wokingham, UK: Addison-Wesley.

Quinonez, D. (1998). Common applications of electrophysiology (EEG) in the past and today: the technologist's view. *Electroencephalography and clinical neurophysiology, 106*(2), 108-112.

Ramos, E., James, C., & Bear-Lehman, J. (2005). Children's computer usage: are they at risk of developing repetitive strain injury. *Medline, 25*(2), 143-154.

Rautaray, S. S., & Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review, 1*(43), 1–54.

Reimer, J. (2005). A History of the GUI. *ARS Technica, 5*.

Rekha, J., Bhattacharya, J., & Majumder, S. (2011). Shape, Texture and Local Movement Hand Gesture Features for Indian Sign Language. *In the Proceedings of the 3rd International Conference on Trendz in Information Sciences and Computing (TISC)* (pp. 30-35). IEEE.

Reyes, J. F., & Tosunoglu, S. (2011). An overview of Brain-Computer Interface Technology Applications in Robotics. *In the Proceedings of the Florida Conference on Recent Advances in Robotics.* Gainesville.

Ronkko, J., Markkanen, J., Launonen, R., Ferrino, N., Gaia, E., Basso, V., Laukkanen, S. (2006). Multimodal astronaut virtual training prototype. *International Journal of Human-Computer Studies, 64*(3), 182–191.

Rosas-Cholula, G., Ramirez-Cortes, J. M., Alarcon-Aquino, V., Gomez-Gil, P., Rangel-Magdaleno, J. D., & Reyes-Garcia, C. (2013). Gyroscope-Driven Mouse Pointer with an EMOTIV® EEG Headset and Data Analysis Based on Empirical Mode Decomposition. *Sensors, 13*(8), 10561-10583.

Rubin, J., & Chisnell, D. (2008). *Handbook of Usability Testing: Howto Plan, Design, and Conduct Effective Tests* (2nd ed.). John Wiley & Sons.

Saponas, T. S., Tan, D. S., Morris, D., Balakrishnan, R., Turner, J., & Landay, J. A. (2009). Enabling Always-Available Input with Muscle-Computer Interfaces. *In Proceedings of the 22nd annual ACM symposium on User interface software and technology* (pp. 167-176). Victoria, British Columbia, Canada: ACM.

Sauro, J., & Kindlund, E. (2005). A Method to Standardize Usability Metrics into a single score. *In the Proceedings of the SIGCHI conference on Human factors in computing systems ACM* (pp. 401-409). ACM.

Savva, N., & Bianchi-Berthouze, N. (2012). Automatic recognition of affective body movement in a video game scenario. In *Intelligent Technologies for Interactive Entertainment* (pp. 149-159). Berlin Heidelberg: Springer .

Schabowsky, C. N., Godfrey, S. B., Holley, R. J., & Lum, P. S. (2010). Development and pilot testing of HEXORR: hand EXOskeleton rehabilitation robot. *Journal of Neuro-engineering and Rehabilitation, 7*(36), 1-16.

Sesto Senzo. (2012). *Sesto Senzo*. Retrieved 04 22, 2015, from http://www.sestosenso3d.com/home/it/carrello.html?page=shop.browse&category_id=20

Shackel, B. (1991). Usability-context, framework, definition, design and evaluation. In *Human factors for informatics usability* (pp. 21-37). Cambridge University Press.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Wadsworth : Cengage learning.

Shahabi, C., Chung, Y. Y., Huang, K. C., Khoshgozaran-Haghighi, J., Knoblock, C., Lee, S. C., Thakkar, S. (2006). Geodec: Enabling geospatial decision making. *In the Proceedings of the IEEE International Conference and Expo on Multimedia* (pp. 93–96). IEEE.

Sharma, V. K. (1975). An easy method of constructing Latin square designs balanced for the immediate residual and other order effects. *Canadian Journal of Statistics, 2*(1), 119-224.

Shin, J., & Hong, K. (2005). An Improved HCI Method and Information Input Device Using Gloves for Wearable Computers. In the Proceedings of the 10th Panhellenic conference on Advances in Informatics.

Shneiderman, B. (1992). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (2nd ed.). Reading, MA: Addison-Wesley.

Singer, M. (2008). Brain Games. *Technology Review, 111*(4), 82-84.

Sporka, A. J., Kurniawan, S. H., Mahmud, M., & Slavik, P. (2006). Non-speech input and speech recognition for real-time control of computer games. *In the proceedings of the 8th International ACM Conference on Computers and Accessibility*, (pp. 213-220).

Steffan, R., Schull, U., & Kuhlen, T. (1998). Integration of virtual reality based assembly simulation into CAD/CAM environments. *In the Proceedings of the 24th Annual Conference of the IEEE in Industrial Electronics Society. 4*, pp. 2535–2537. IEEE.

Stehman, S. V. (1999). Basic probability sampling designs for thematic map accuracy assessment. *International Journal of Remote Sensing, 20*(12), 2423-2441.

Steinberg, G. (2012). *Natural user interfaces.* Austin: ACM.

Stufflebeam, R. (2008). Neurons, Synapses, Action Potentials, and Neurotransmission. *In Proceedings of the Consortium on Cognitive Science Instruction.*

Sturman, D., & Zeltzer, D. (1994). A survey of glove-based input. *IEEE Computer Graphics and Applications, 14*(1), 30-39.

Swanson, N. G., Galinsky, T. L., Cole, L. L., Pan, C. S., & Sauter, S. L. (1997). The impact of keyboard design on comfort and productivity in a text-entry task. *Applied Ergonomics, 28*(1), 9-16.

Swenson, R. S. (2006). *Review of Clinical and Functional Neuroscience.* Dartmouth : Dartmouth Medical School.

Tan, D. S., & Nijholt, A. (2010). *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction.* London: Springer.

Tangermann, M., Krauledat, M., Grzeska, K., Sagebaum, M., Blankertz, B., Vidaurre, C., & Müller, K. R. (2009). Playing Pinball with non-invasive BCI. *Advances in Neural Information Processing Systems*(21), 1641-1648.

Tani, B., Maia, R. F., & von Wangenheim, A. (2007). A gesture interface for radiological workstations. *In the Proceedings of the 20th IEEE International Symposium on Computer-Based Medical Systems* (pp. 27–32). IEEE.

Tansey, O. (2007). Process Tracing and Elite Interviewing: A Case for Non-Probability Sampling. *Political Science and Politics, 40*(4), 765-772.

Taveira, A. D., & Choi, S. D. (2009). Review Study of Computer Input Devices. *International Journal of Human-computer Interaction, 25*(5), 455–474.

Tonin, L., Leeb, R., Tavella, M., Perdikis, S., & del Millan, J. R. (2010). The role of shared-control in BCI-based telepresence. *In the Proceedings of the Systems Man and Cybernetics 2010 IEEE International Conference* (pp. 1462-1466). IEEE.

Trochim, W. (2005). *Research methods: The concise knowledge base.* Atomic Dog Publishing. Retrieved 08 20, 2013, from http://www.human.cornell.edu

Tudor, M., Tudor, L., & Tudor, K. (2005). Hans Berger (1873-1941) - the history of electroencephalography. *Acta Med Croatica, 59*(4), 307-313.

Tufte, E. R. (1989). *Visual Design of the User Interface.* Armonk NY: IBM Corporation.

Tullis, T., & Albert, B. (2013). *Measuring the User Experience* (2nd ed.). Waltham: Elsevier Inc.

Turk, M. (2014). Multimodal interaction: A review. *Pattern Recognition Letters, 36*, 189-195.

Usakli, A. B., Gurkan, S., Aloise, F., Vecchiato, G., & Babiloni, F. (2010). On the Use of Electrooculogram for Efficient Human Computer Interfaces. *Computational Intelligence and Neuroscience 2010, 2010*(1), 135629.

Utting, M., & Legeard, B. (2010). *Practical Model-Based Testing: a Tools Approach.* Morgan Kaufmann.

Vallabhaneni, A., Wang, T., & He, B. (2005). Brain Computer Interface. In *Neural Engineering* (pp. 85-121). Minnesota: Springer.

van de Laar, B., Plass-Oude Bos, B. R., Reuderink, B., Poel, M., & Nijholdt, A. (2013). How Much Control Is Enough? Influence of Unreliable Input on User Experience. *IEEE Transactions on Cybernetics, 6*(43), 1584-1592.

van Erp, J. B., Lotte, F., & Tangermann, M. (2012). Brain-Computer Interfaces: Beyond Medical Applications. *Computer, 45*(4), 26-34.

Vasantrao, S. S., Prakash, C. M., Prakash, B. A., & Anant, G. A. (2014). Improved HCI using face detection and speech recognition. *International Journal, 3*(2).

Vasey, M. W., & Thayer, J. F. (1987). The continuing problem of false positives in repeated measures ANOVA in psychophysiology: A multivariate solution. *Psychophysiology, 24*(4), 479-486.

Virtual Realities Ltd. (2013). *Virtual Realities*. Retrieved 04 22, 2015, from https://www.vrealities.com/products/data-gloves/dg5-glove-3-0

Volosyak, I., Cecotti, H., & Graser, A. (2009). Impact of frequency selection on LCD screens for SSVEP based brain-computer interfaces. In *Bio-Inspired Systems: Computational and Ambient Intelligence* (pp. 706-713). Berlin Heidelberg: Springer.

Wigdor, D., & Wixon, D. (2011). *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture.* Elsevier.

Wolpaw, J. R. (2014). The BCI endeavor and the mission of this new journal. *Brain-Computer Interfaces, 1*(1), 2-4.

Wolpaw, J. R., & McFarland, D. J. (2004). Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, pp. 17849-17854.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology, 113*(6), 767-791.

Wolpaw, J. R., Loeb, G. E., Allison, B. Z., Donchin, E., do Nascimento, O. F., Heetderks, W. J., Turner, J. N. (2006). BCI Meeting 2005—Workshop on Signals and Recording Methods. *Neural Systems and Rehabilitation Engineering, 14*(2), 138-141.

Wolpaw, R., McFarland, D., Neat, G., & Forneris, C. (1991). An EEG-based brain-computer interface for cursor control. *Electroencephalography and Clinical Neurophysiology, 78*(3), 252-259.

Xu, D., Yao, W., & Zhang, Y. (2006). Hand gesture interaction for virtual training of SPG. *In the Proceedings of the 16th International Conference on in Artificial Reality and Telexistence* (pp. 672–676). IEEE.

Appendix A

Head mounted accelerometer + Facial Expressions as well as

Head mounted accelerometer + Peregrine glove

Questionnaire for User Interface Satisfaction (QUIS).

QUIS

Please tick the circle that is most appropriate as an answer to the given comment.

Overall reaction to the user interface

	1	2	3	4	5	6	7	8	9	
Terrible	○	○	○	○	○	○	○	○	○	Wonderful
Difficult	○	○	○	○	○	○	○	○	○	Easy
Frustrating	○	○	○	○	○	○	○	○	○	Satisfying
Inadequate power	○	○	○	○	○	○	○	○	○	Adequate power
Dull	○	○	○	○	○	○	○	○	○	Stimulating
Rigid	○	○	○	○	○	○	○	○	○	Flexible

Appendix B

Pointing device assessment questionnaire

Device assessment
Please circle the number that is most appropriate as an answer to the given comment.

1. The force required for actuation (propelling or moving the device) was

1	2	3	4	5
too low				too high

2. Smoothness during operation was

1	2	3	4	5
very rough				very smooth

3. The mental effort required for operation was

1	2	3	4	5
too low				too high

4. The physical effort required for operation was

1	2	3	4	5
too low				too high

5. Accurate pointing was

1	2	3	4	5
easy				difficult

6. Operation speed was

1	2	3	4	5
too fast				too slow

7. Neck fatigue

1	2	3	4	5
none				very high

8. General comfort:

1	2	3	4	5
very uncomfortable				very comfortable

9. Overall, the input device was

1	2	3	4	5
very difficult to use				very easy to use

Any other comments and suggestions:

……
……
……
……
……
……

Appendix C

Balanced Latin Square Task and User Interface randomisation

Task 1 = T_1 User Interface 1 = UI_1

Task 2 = T_2 User Interface 2 = UI_2

Task 3 = T_3 User Interface 3 = UI_3

1 – T_1UI_1 **4 – T_2UI_1** **7 – T_3UI_1**

2 – T_1UI_2 **5 – T_2UI_2** **8 – T_3UI_2**

3– T_1UI_3 **6 – T_2UI_3** **9 – T_3UI_3**

Formula = 1, 2, n, 3, n-1, 4, n-2, 5, n-3

Subjects	1ST	2ND	3RD	4TH	5TH	6th	7th	8th	9th
A	1	2	9	3	8	4	7	5	6
B	2	3	1	4	9	5	8	6	7
C	3	4	2	5	1	6	9	7	8
D	4	5	3	6	2	7	1	8	9
E	5	6	4	7	3	8	2	9	1
F	6	7	5	8	4	9	3	1	2
G	7	8	6	9	5	1	4	2	3
H	8	9	7	1	6	2	5	3	4
I	9	1	8	2	7	3	6	4	5
J	6	5	7	4	8	3	9	2	1
K	7	6	8	5	9	4	1	3	2
L	8	7	9	6	1	5	2	4	3
M	9	8	1	7	2	6	3	5	4
N	1	9	2	8	3	7	4	6	5
O	2	1	3	9	4	8	5	7	6
P	3	2	4	1	5	9	6	8	7
Q	4	3	5	2	6	1	7	9	8
R	5	4	6	3	7	2	8	1	9

Appendix D

Test Protocol

During the usability study, the following protocol will be followed for each session with a participant:

1. Participant will be welcomed into the usability laboratory. He/she will be seated and then asked to complete the consent form.
2. After completing the consent form, the goal of the study and the testing routine will be explained in detail. Special attention will be given to how the BCI and glove technology should be used.
3. The participant will subsequently be asked to complete the pre-test questionnaire (as seen in Appendix A) which also requires the participant to enter his/her biographical details.
4. If it is the first session, a quick demonstration of the BCI and Peregrine gaming glove is conducted with the participant. During this demonstration the participant is allowed and encouraged to ask questions on the functioning of the devices and any unclear parts which need more explanation.
5. Each participants is also allocated, as a subject, to the Latin square (as seen in Appendix D).
6. After this initial introduction and gathering of details, the participant will be prepared for the study. This will include aspects like fitting the BCI headset or the Peregrine gaming glove.
7. The following step is to start with the actual testing of the participant. If it is the first session, this includes a training task to familiarise the participant with the interface. Subsequent sessions will then be completed without any training, as the participant will then be familiar with the interface at hand. After the first session the participant will be asked to complete a post-task questionnaire that will also be done after the completion of the final session. A post-test questionnaire will be completed for each of the alternative user interface combination per participant during these mentioned sessions. The steps that will be followed for the actual testing of each participant are as follows:
    a. Before the participant may start with the first task he/she will be provided with a training task in order to familiarize him/her with the current user interface. For the training task Microsoft Word will be used to check whether the participant is able to activate the two actions required to play the game (Reload and Shoot).
    b. The researcher will start the selected task and the participant will be presented with the selected interface as per experimental setup.

c. The participant will complete the task. After completion the researcher will allow for 2 minutes of rest and then start the following task and provide the participant with the following user interface.

8. When all nine parameter combinations have been tested, the researcher will close the game and retrieve the user interfaces.
9. The participant will then be asked to complete the post-test questionnaire. After completion of this questionnaire the researcher will thank the participant for his/her participation in the study and provide the participants with an incentive in the form of a voucher.
10. This concludes the session.

Appendix E

Natural User Interface study pre-test questionnaire

University of the Free State

Department of Computer Science and Informatics

Pre-test Questionnaire

ON BEHALF OF THE UNIVERSITY OF THE FREE STATE AND THE DEPARTMENT OF COMPUTER SCIENCE WE WOULD LIKE TO THANK YOU FOR PARTICIPATING IN THIS RESEARCH PROJECT.

WE OFFER OUR ASSURANCE THAT ALL INFORMATION RECORDED HERE WILL ONLY BE USED FOR RESEARCH PURPOSES AND YOUR PARTICIPATION IS VOLUNTARY. PLEASE ANSWER THE REQUIRED QUESTIONS AND WHEN NECESSARY INDICATE THE APPROPRIATE ANSWER WITH AN X.

PLEASE ANSWER THE FOLLOWING QUESTIONS:

1. Participant ID (***will be provided by the facilitator***): ……

2. Age: ……..

3. Gender: Male / Female

4. Home Language: ……

5. For how many years have you been using a computer?

☐	Never used a computer
☐	Less than 1 year
☐	1 – Less than 3 years
☐	3 – Less than 5 Years
☐	5 years or more

If you answered "Never used a computer" then skip questions 6 – 13.

6. How often do you use a computer?

- [ ] Daily
- [ ] Weekly
- [ ] Once every two weeks
- [ ] Once a month
- [ ] Less than once a month

7. For how many years have you been using a computer mouse?

- [ ] Never used a mouse
- [ ] Less than 1 year
- [ ] 1 – 3 less than 3 years
- [ ] 3 - less than 5 Years
- [ ] 5 years or more

If you answered "Never used a mouse" then skip question 8 and proceed to question 9.

8. How often do you use a computer mouse?

- [ ] Daily
- [ ] Weekly
- [ ] Once every two weeks
- [ ] Once a month
- [ ] Less than once a month

9. Do you play computer games?

- [ ] Yes
- [ ] No

If yes, which games do you regularly play?

……

……..

If Yes, proceed to Question 10, else proceed to Question 11.

10. How often do you play computer games?

- [ ] Daily
- [ ] Weekly
- [ ] Once every two weeks
- [ ] Once a month
- [ ] Less than once a month

11. Have you ever used a Brain Computer Interface to work on a computer?

- [ ] Yes
- [ ] No

If Yes, proceed to Question 12, else proceed to Question 13.

12. Have you ever used a Brain Computer Interface as a pointing device (substitute for a mouse)?

- [ ] Yes
- [ ] No

13. Have you ever used a Glove-based input device to work on a computer?

- [ ] Yes
- [ ] No

Appendix F

Ethical Clearance

Faculty of Natural and Agricultural Sciences

23-Oct-2014

Dear Mr Rouxan Fouche

Ethics Clearance: INVESTIGATING THE USABILITY OF NATURAL USER INTERFACES FOR GAMEPLAY

Study Leader/Supervisor: Beelders, Tanya

Principal Investigator: Mr Rouxan Fouche

Department: Computer Science and Informatics (Bloemfontein Campus)

This letter confirms that a research proposal with tracking number: **UFS-HSD2014/0333** and title: **'INVESTIGATING THE USABILITY OF NATURAL USER INTERFACES FOR GAMEPLAY'** was given ethics clearance by the Ethical Committee.

Please ensure that the Ethical Committee is notified should any substantive change(s) be made, for whatever reason, during the research process. This includes changes in investigators. Please also ensure that a brief report is submitted to the Ethical Committee on completion of the research. The purpose of this report is to indicate whether or not the research was conducted successfully, if any aspects could not be completed, or if any problems arose that the Ethical Committee should be aware of.

Note:
1. This clearance is valid from the date on this letter to the time of completion of data collection.
2. Progress reports should be submitted annually unless otherwise specified.

Yours Sincerely

Prof. Neil Heideman
Chairperson: Ethical Committee
Faculty of Natural and Agricultural Sciences

Abstract

This study aimed to determine to what extent the usability of a two dimensional game was influenced by the use of a Natural User Interfaces (NUI) as opposed to a traditional keyboard and mouse combination. Two multimodal NUIs were investigated during the study. The first NUI combination (BCIG) made use of the Peregrine gaming glove for the activation of commands, combined with the Emotiv's accelerometer for control of the cursor. The second NUI combination (BCIF) made use of facial expression recognition, offered by the Emotiv Brain Computer Interface (BCI), as a method of command activation in combination with the Emotiv's built-in accelerometer for cursor control. A shooting genre game was developed and three tasks were included during development to simulate gaming actions. The first task used only stationary targets, the second task used predictable moving targets, whereas the third task made use of unpredictable moving targets. Since the Emotiv BCI allows for customisation of the accelerometer sensitivity settings, a pilot study was conducted to determine whether the low, medium or high sensitivity setting would provide the best cursor control. The low sensitivity resulted in the fastest gameplay overall as well as the least number of errors. It could thus be concluded that the lowest setting is the optimal setting since it provided the most efficient control for three out of the four metrics tested. After implementing this result, user testing, which involved 5 sessions per participant (n=18), was conducted. Data for three metrics was gathered during user testing, which included data on effectiveness, efficiency and learnability. Post-test questionnaires were administered to assess the level of user satisfaction with each NUI. The results of this study indicated that there is a difference between the usability of the traditional input combination, the keyboard and mouse, and the two NUIs investigated in this study. With regard to the effectiveness and efficiency metrics the traditional input combination was found to be the best, closely followed by BCIG. The three interfaces showed dissimilar levels of improvement, with the traditional keyboard and mouse combination showing the least, followed by BCIG, while the best improvement was noticed for BCIF. By analysing the subjective data gathered from post-test questionnaires, it was found that participants initially enjoyed using BCIG, and after several sessions their level of satisfaction improved. In comparison, the participants initially experienced a slightly negative feeling towards BCIF, which then improved over several sessions to a positive overall response. In conclusion, the keyboard and mouse combination provided far more effective and efficient input, with one exception being the fast command activation when making use of the Peregrine glove, where the two interfaces compared well.

It was found that a significant obstacle in the way of NUIs is the existing skill and acceptance that computer users have with the traditional interface combination. Consequently, for individuals to accept and migrate to a more natural interface the new interface will have to provide more effective and efficient input than what is already achievable with the keyboard and mouse combination.

Opsomming

Hierdie studie het gepoog om te bepaal tot watter mate die bruikbaarheid van 'n twee-dimensionele speletjie beïnvloed word deur die gebruik van 'n natuurlike gebruikerskoppelvlak (NUI) in teenstelling met 'n tradisionele sleutelbord-en-muis kombinasie. Tydens die studie is twee multimodale NUIs ondersoek. Die eerste NUI kombinasie (BCIG) het van die Peregrine-handskoen gebruik gemaak vir die aktivering van opdragte in kombinasie met die Emotiv se versnellingsmeter ("accelerometer") vir die beheer van die posisiemerker. Die tweede NUI kombinasie (BCIF) het gebruik gemaak van die Emotiv Brein-rekenaarkoppelvlak (BCI) se gesigsuitdrukking-herkenning funksie vir opdrag-aktivering in kombinasie met die versnellingsmeter vir posisiemerker beheer. 'n Skiet-genre speletjie is ontwikkel en drie take is ingesluit tydens die ontwikkeling om die aksies van rekenaarspeletjies na te boots. Die eerste taak maak gebruik van stilstaande teikens, die tweede taak van voorspelbare bewegende teikens, terwyl die derde taak gebruik maak van onvoorspelbare bewegende teikens. Aangesien die Emotiv BCI toelaat vir die verstelling van die versnellingsmeter se vlakke van sensitiwiteit is 'n loods-studie gedoen om te bepaal of die lae, medium of hoë sensitiwiteit stelling die beste muis-beheer sal voorsien. Die lae sensitiwiteit was verantwoordelik vir die algehele vinnigste spel sowel as die minste aantal foute. Dit kan dus afgelei word dat die laagste stelling die beste stelling is aangesien dit die doeltreffendste beheer vir drie van die vier metings verskaf het. Na die implementering van die resultaat is gebruikerstoetse gedoen, waar 5 sessies per deelnemer (n = 18) voltooi is. Data vir drie metings is versamel tydens die gebruikerstoetse, insluitende data ten opsigte van effektiwiteit, doeltreffendheid en leerbaarheid. Post-toets vraelyste is voltooi om die gebruikers se vlak van tevredenheid ten opsigte van die NUI te evalueer. Die resultate van hierdie studie het aangedui dat daar 'n verskil tussen die bruikbaarheid van die tradisionele kombinasie, die sleutelbord en muis, en die twee NUIs voorkom. Daar is gevind dat die tradisionele kombinasie gevolg deur BCIG die beste is met betrekking tot die effektiwiteit- en doeltreffendheidmetings. Die drie koppelvlakke het uiteenlopende vlakke van verbetering getoon. Die tradisionele sleutelbord-en-muis kombinasie het die minste verbetering getoon, gevolg deur BCIG, terwyl die beste verbetering opgemerk is vir BCIF. Deur die ontleding van die subjektiewe data wat uit post-toets vraelyste verkry is, is daar gevind dat die deelnemers aanvanklik die BCIG geniet het, en na 'n paar sessies het hul vlak van tevredenheid verbeter. In vergelyking hiermee het die deelnemers aanvanklik 'n effens negatiewe gevoel teenoor BCIF ervaar, wat oor verskeie sessies verbeter het tot 'n algehele,

positiewe reaksie. Ten slotte, die sleutelbord-en-muis kombinasie verskaf veel meer effektiewe en doeltreffende beheer, met een uitsondering, naamlik die Peregrine-handskoen se vinnige opdrag-aktivering, waar die twee koppelvlakke goed vergelyk het. Daar is gevind dat die bestaande vaardigheid en aanvaarding van rekenaar-gebruikers met die tradisionele koppelvlak-kombinasie 'n groot struikelblok in die weg van NUIs is. Gevolglik moet meer effektiewe en doeltreffende beheer as dit wat reeds bereikbaar is met die sleutelbord-en-muis kombinasie gelewer word vir individue om meer natuurlike koppelvlakke te aanvaar en te gebruik.