# Evaluating HCI using the Microsoft Kinect augmented with Non-invasive BCI

By Peter Mitchell

**Qualification: Bachelor of Computer Science**

**November 2012**

**Supervised by Dr. Brett Wilkinson**

Submitted to the School of Computer Science, Engineering, and Mathematics in the Faculty of Science and Engineering in partial fulfilment of the requirements for the degree of Bachelor of Computer Science at Flinders University – Adelaide Australia.

# Declaration

I certify that this work does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signature: _____________________ Date: ________________

## Abstract

Natural User Interfaces (NUI) enable a varied scope of what can be used when it comes to computer input. The standard paradigm of a mouse or keyboard limits the interaction that can be completed through exclusion of inputs related to more natural physical movement, speech, or even thought. Brain Computer Interfaces (BCI) and the Microsoft Kinect are two examples of technology that can be harnessed to provide NUI. The purpose of this research was to investigate whether BCI can augment the gameplay experience provided by the Microsoft Kinect. Gürök et al. (2012) wrote that many of the BCI games being developed focus on "psychological hypotheses or evaluating the performance of signal analysis and classification techniques". Gürök et al. go on to say that "experience evaluations are almost never carried out".

To accomplish the goal of determining how viable the combinations of inputs are, three computer game puzzles were created and then assessed through testing and a basic user experience evaluation. Aspects of these games could be controlled by the user by altering their levels of concentration and relaxation. These levels were determined from an electroencephalography (EEG) reading of Alpha waves captured by an Emotiv EPOC (a BCI device). The variation in alpha levels provided gameplay augmentation, while the primary game inputs were provided by the Microsoft Kinect using speech and skeletal tracking. Participants who took part in this evaluation completed a calibration stage to prepare the system for their specific alpha variations. The games implemented included a tile puzzle which contained a hidden image that could be revealed through BCI interaction and then correctly ordered using gestures. The other two games provided more active puzzles involving following a target using the BCI to modify speed of travel, and a river puzzle that involved slowing down the perceived time to make it easier to grab objects in the river.

There were 15 participants who took part in the experiment. It was found that there was a high level of challenge particularly in changing mental states. It was suggested there was replay-ability and a medium level of entertainment. In its current form the ease of use was rated poorly particularly as 40% of individuals failed to complete the first puzzle which respondents suggested was largely due to the inaccuracies of the voice recognition. Limitations of the BCI system were also identified with muscle movement accounted for a very high proportion of the time spent in an apparently relaxed state. In the future it is foreseen that further testing will take place with more complex games and the system will be developed to explore this type of technology's potential in other areas of research.

# Table of Contents

## Table of Figures

## Table of Tables

## Acknowledgements

I would like to acknowledge a few individuals who have contributed to the development of this project. I would like to acknowledge the work of Dr. Brett Wilkinson as my supervisor for providing input throughout the project when needed. I would also like to acknowledge the advice and consultation provided by Dr. Sean Fitzgibbon in relation to the collaborative development of the server side application for extracting the data from the Emotiv EPOC and other general BCI instruction. I would like to acknowledge the time of those who spent time participating in the pilot study for this research. I would like to acknowledge the Flinders University Social and Behavioural Research Ethics Committee who provided Ethics approval to conduct the experiment (Project No. 5633). Finally I would like to acknowledge the support provided by family and friends.

# 1 Introduction

## 1.1 Project Overview

Natural User Interaction (NUI) technologies allow for engagement with computing devices without the tether of mouse or keyboard; Brain computer interfaces (BCI) and the Microsoft Kinect are examples of new devices that support this type of user experience. This project investigated whether BCI can augment the gameplay experience provided by the Microsoft Kinect (Figure 1). The Emotiv EPOC (Figure 2) was used as the specific hardware solution for the BCI input as it is available for purchase as a consumer level product with adequate reliability for the scenario employed by this research project. The use of this product is commercially more attractive for future applications because of the relative cheap price when compared to the high cost medical-grade BCI equipment. The Emotiv EPOC provides a simple EEG reading[1]. The reading from the Emotiv EPOC device specifically is used to determine the state of concentration or relaxation. It is important to recognise the BCI element of this research was intended to always be simply an augmentation. For this project the BCI element has been where the most difficulty of product design and development has occurred. The intention was to create a system that works effectively with this input in a simple manner. The other input technologies provided by the Microsoft Kinect enable a full body experience. The BCI provides access to the brain and emotional states, while the Microsoft Kinect provides skeletal tracking and speech recognition.


Figure 1. Microsoft Kinect

Figure 2. Emotiv EPOC

To accomplish the goals of the project a simple experiment was designed. The three stages in the experiment included: a calibration stage, completion of three different puzzles, and a short questionnaire to provide feedback about the experience. The calibration stage allows for individual participant variations in brain wave values to be presented in appropriate value ranges for the algorithms used to process the BCI elements of gameplay. The three puzzles were a Tile Puzzle, Street Puzzle, and River Puzzle. Each provided testing of a different level of interaction using the different inputs. The questionnaire was used to establish relevant background information about participants and their thoughts about the experience using the hardware.

There were 15 participants who took part in the experiments. Ethical approval to conduct the experiment with volunteers was sought from and granted by Flinders University Social and Behavioural Research Ethics Committee (Project No. 5633). The experiment may be considered a

[1] The use of existing BCI software developed by Dr Sean Fitzgibbon was utilised by this project.

pilot study for this combination of inputs with this number of participants. It is suspected the results have a high contamination as a result of muscle movement due to the limitations of EEG equipment used. It was found that on average 71% of the time was spent in a relaxed state and 29% in a concentrating state. Muscle movement registers as if the subject were in a relaxed state, which makes this statistic make more sense, as it seems unreasonable that participants could have spent so much time relaxed throughout the playing of the puzzles. From the questionnaire it was determined that the puzzles were challenging while difficult in regard to their ease of use, with a reasonable level of replay-ability and entertainment. Future studies should ideally test with a larger group of participants and trial use of the systems for multiple sessions to gauge the learning curve that is experienced. There are other areas of future work that will be discussed at the end of this document.

## 1.2 Research Goals

The overall research question used to determine the goals of this project was: "To what extent is the use of BCI a viable extension to the HCI input provided by the Microsoft Kinect?" To answer this question a number of more specific sub-goals were identified:

- Does the inclusion of BCI:
    - improve the entertainment value of gameplay?
    - provide a sufficient extra level of challenge?
    - provide replay-ability?
- How difficult is the ease of use for the three input technologies (BCI, gesture, speech) when considered separately and together?
- What sorts of strategies might be employed for dealing with the input technologies in a simplistic manner that could be adopted for use in projects in the future?

## 1.3 Contents of the document

The remaining sections of this document will go into detail about the specific aspects of this research project. The literature review covers relevant research that has been reviewed. The experiment design is divided into two separate sections. The first provides an overview of the experiment identifying details specific to the individual puzzles and other gameplay related elements. The second section covers the software design discussing an overview of the important aspects that were considered - specifically the detail of how the Microsoft Kinect and Emotiv EPOC were integrated into the experiments. The results and results analysis sections provide an overview of the questionnaire and quantitative data collected by the application. Also the analysis section provides identification of the issues with the technologies that have been either worked around or must be revised to lead to future work. The future work and conclusions wrap up the content of this research providing expansion for the future of this research.

## 2 Literature Review

In recent years there have been many studies investigating the usability of brain computer interaction systems. Some of these have been focussed on interactions in a video game environment. Many of these have used the input as the core element of the gameplay such as BrainBasher (Van de Laar et al., 2008). In BrainBasher (Figure 3 contains a screenshot of the game) the player was required to complete actions by using imagined and physical movement of the hands. They would score points based on how quickly and accurately the actions were performed.

The following quote is from a paper by Nijholt et al. It emphasises the relevance of providing new experiences that can revolutionise not only gaming, but expand into other fields.

> *"Games are usually the first applications to adopt new paradigms, driven by gamers continuing search for novelty and challenges."* (Nijholt et al., 2008)

Providing natural interfaces that allow the control to be interpreted by the computer is one way that novelty and challenge can be included. The idea of using natural interfaces is one among many other attributes the paper by Nielsen (1993) suggested as the future for computing. To extend this a step further, a full body experience is proposed. Despite the variety of sample tests using brain computer interfaces in the area of games development there has not yet been any identified work that combines the Microsoft Kinect or similar devices. By combining a device such as the Microsoft Kinect the amount of interaction that can be provided expands.

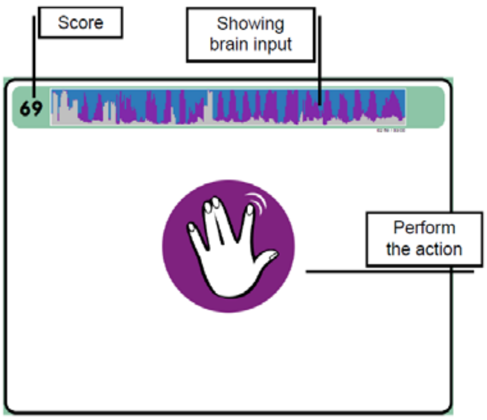



Figure 3. BrainBasher (Van de Laar, 2008)

Games such as BrainBasher require a low acceptance threshold for the comparison between actual and imagined movement to account for the variation between players. 60% was identified as a threshold that was used successfully in the example of BrainBasher (Van de Laar et al., 2008). The brain is a complicated system to interpret for individual activities, even in the case of simple tasks. BrainBasher has been used by Van de Laar et al. to compare the use of actual movement against imagined movement. To accomplish this comparison BrainBasher required a calibration stage using machine learning techniques to train the computer. The neurological process used in this paper was referred to as Event-Related Desynchronization. This approach uses the beta frequencies to match against movement. The issue of participants finding they were incapable of completing the task was a noted result of this research. This lack of being able to be understood by the BCI was referred to from other research as BCI illiteracy.

Other studies have taken a similar approach to those in BrainBasher using the brain as the primary method of control. An application called AlphaWoW by Bos et al. (2010) was one of a few applications created by these researchers to test the combination of BCI, and human computer interaction (HCI). Approaches such as “inner speech” (reciting a mental spell), “association” (feeling what you want to be), and mental states (emotional fear used as a trigger) were used to change the player’s character between specified in-game forms. The application was using the game World of Warcraft and simply allowing the changing between a bear form and a humanoid form. The human form was intended to leave the player feeling more vulnerable. Therefore the bear form was used to signify a form of protection that could be activated. The mental state of fear used as the trigger to activate the bear form, which would then remain active until the fear had reduced. Another part of the study used an application called “Bacteria Hunt” (Bos et al., 2010) (Figure 4 contains a screenshot of the game) that utilised the player’s alpha state as a speed modifier to control movement in combination with the keyboard. Alpha state refers to a specific frequency range of rhythmic waves that can be detected with a BCI device (Niedermeyer and Da Silva 2004a). They relate to the relaxation or activity of an individual, and will be discussed further with discussion of the Emotiv EPOC BCI device.

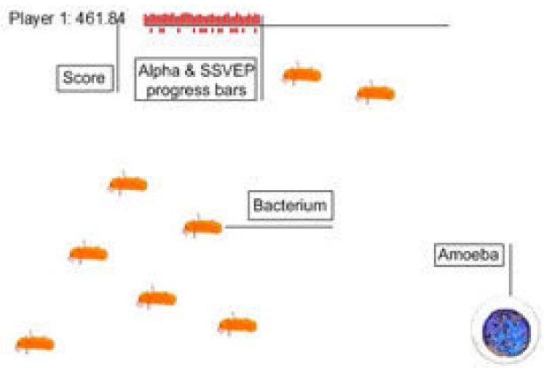



**Figure 4. Bacteria Hunt (Bos, 2010)**

It is important that devices used to research game interaction are cheap enough to be potentially marketable in the future. If no one can afford to buy the equipment needed the point of researching new technologies using the area of games is diminished. The Microsoft Kinect has already entered the market, and there are brain devices that are also readily available for use. A couple of these are the MindFlex Duel (ThinkGeek, 2012) a concentration based two player game, and the Neurosky (Neurosky, 2012) a more open ended platform intended to detect basic states of the brain. For this project the Emotiv EPOC device has been selected as the BCI device. It comes with a larger number of sensors than the other devices providing the potential for improved responsiveness. The device will not be used as the primary means of motor based movement, studies by Cutrell and Tan (2004) suggested that this was not something that would be adopted by the general population. As the price of these forms of hardware is reduced it may become a more commonly accepted input in an evolving society.

The Emotiv EPOC device is designed with the purpose of detecting emotional states. These states are determined by using a variety of different rhythmic waves such as alpha waves that the brain produces (Niedermeyer and Da Silva, 2004a; b; c). The extract below discusses the properties of the alpha wave:

> *"Rhythm at 8–13 Hz occurring during wakefulness over the posterior regions of the head, generally with higher voltage over the occipital areas. Amplitude is variable but is mostly below 50 μV in adults. Best seen with eyes closed and under conditions of physical relaxation and relative mental inactivity. Blocked or attenuated by attention, especially visual, and mental effort (IFSECN, 1974)."* (Niedermeyer and Da Silva, 2004c)

It is suggested that relaxation and concentration emotional states can be used to augment the gameplay experience. Additional brainwaves that could benefit the experiment could include beta and perhaps gamma waves. This type of emotional BCI has been researched by Molina et al. (2009) who determined a subset of emotions that can be used to derive any other emotion. The area of BCI is continually developing new methods and devices that improve on those that are available, with current research into areas like controlling virtual characters in virtual environments using just BCI input (Lecuyer et al., 2008).

Other BCI research includes the work of Wolpaw et al. (2006) who discusses BCI as a means of assisting the movement of impaired limbs, Schalk et al. (2004) have designed system for managing BCI systems, and Ebrahimi et al. (2003) consider BCI to be an expression that is comparable to how audio/visual/taste expressions have been used in the past. Significantly van Erp et al. (2011) in a yet to be released paper have suggested 7 non-medical applications that BCI will expand into in the coming years. The paper suggests that the area that should and will be the first to uptake BCI will be gaming followed by other areas such as training and education. Training is identified in other papers too as an area for further development (Graimann et al., 2007). The following identifies the full list presented by Erp et al. and the time frames suggested for when the systems will arrive on the market using BCI as an input:

- Gaming and Entertainment (now)
- Evaluation (1 to 3 years)
- Cognitive Improvement (3 to 5 years)
- Training and Education (3 to 5 years)
- User State Monitoring (3 to 5 years)
- Control of Devices (5 to 10 years)
- Safety and Security (5 to 10 years)

Usability, hardware availability, signal processing, and system integration are all stated in the paper as hurdles that must be overcome for these areas of development to occur.

Interesting developments are coming to light which may alter the course of BCI commercial use in the area of security. There has already been some development in this area using the Emotiv EPOC. Martinovic et al. (2012) used the Emotiv EPOC to attempt to determine secret information from individuals. The research was not successful in this attempt based on their trials; however, it may be the first step in other potential security based uses for BCI.

Games can also be developed for medical purposes that integrate BCI as a primary method of control. Wang et al. (2010) have used EEG based gaming to develop concentration driven challenges for the purpose of inducing a healing effect on suffers of Attention Deficit Hyperactivity Disorder (ADHD) and other similar neurological disorders. This is just one instance of BCI used specifically for sufferers of a particular disorder. People who have specific debilitating disorders that reduce their motor movement could find benefit and enjoyment from the "hands free" approach that BCI offers.

Interfaces and HCI methods have been studied at length by many different researchers. The current standard paradigm uses an approach of Windows, Icons, Menus and Pointers (WIMP) as the means for interfacing with computer systems. With the development into areas such as Microsoft Kinect and BCI the standard WIMP interface may be superseded by a wider set of interactions. Gesture interaction and natural language interfaces provide a potentially wider array of options than a mouse and keyboard allow for. Manipulation of shapes using translation, scaling, and rotation are methods that have been used for testing these types of input methods. For example the research of Hauptmann (1989) observed how many fingers or actions participants used to attempt to manipulate graphical objects using the above interactions. This type of simple shape interaction has been partially included as a basis for the game puzzle experiments that will be discussed later in this thesis, for example the first puzzle uses simple squares that require rotation and translation (the Tile Puzzle).

Multi-hand input has also been tested using various types of devices in the work of Jacob (1996). Multi-hand interaction is something that the Microsoft Kinect is well suited for as it tracks the skeleton across 20 specific joints. Although unlike Hauptmann's research (1989) it is not possible to natively detect individual fingers using the Kinect API. Another example where multiple hands have been used for compound tasks is the work of Kabbash et al. (1994). In this research it was found that the best performance came from applying existing skills that would be used in everyday life to the problem solving. It also claims that using a Fitts type analysis (Fitts, 1954) is not sufficient to determine the performance of an input method. The reason given for this is that the "cognitive load must also be considered" (Kabbash et al., 1994). In addition to all this it was discovered that it is possible for two hands to be worse than just one. The commonly cited issue of trying to tap the head while rubbing the stomach was used as an example of a case where two hands are not ideal. Combining the complexity of understanding both the Microsoft Kinect and the Emotiv EPOC device, the use of multiple hands has only been used in one simple instance for the experiment defined in this paper. There is no point in trying to run an experiment if there is not sufficient simplification provided for there to be useful feedback and application outputs. The instance it has been used in this experiment is for stopping the character during an active challenge where the player chases a thief. The raising of both hands above shoulder height has been used to signify a desire by the user to have their character stop.

To match these new input methods papers have been written by the likes of van Dam (1997; 2000; 2001a; 2001b), identifying how the new areas of interface design can be moved into and worked with. Combining the Microsoft Kinect and the Emotiv EPOC device specifically for manipulating interfaces is only one of a variety of combinations that have been attempted. Another example is the use of multi-modal natural dialogue researched by Thorisson et al. (1992). In their research the use of speech, hands, and eye tracking was used as a combined approach to determine how the participants were interacting with the environment. Others such as Bolt and Herranz (1992) who contrasted free-hand with two hands against inputs with a variety of input devices described speech input as “choppy” and gestural data as “spotty”. It is suggested therefore that certainly at the time the complexities of determining relevant details from the useless ones and providing a true multi-modal method of interacting with the computer was not yet ready. A more recent discussion of the evolving state of HCI can be found in a paper by Gentile et al. (2011). The paper included discussion of the evolution from Graphical User Interfaces (GUIs) to Touchless Gesture-based User Interfaces (TGUIs) and Voice User Interfaces (VUIs). These two interface types can be implemented using the Microsoft Kinect. The final stage of HCI discussed in the paper relates to the eventual transition to BCI.

Evaluation of games provides an additional factor of complexity. IJsselsteijn et al. (2007) have suggested that the flow and immersion of games need to be considered in addition to the usability. Thus, as part of the testing process, the participants have been asked to respond to questions that correlate to these types of evaluation methods. This is qualitative in approach as it would be an entire project in itself to define a quantitative approach to properly evaluate games. As cited by Van de Laar et al. (2008) “Gamers are always looking for challenges and limitations that they can overcome by practice.” Based on this reasoning, the chosen method of evaluation for this research is to cover a qualitative view of participants to indicate their views on the challenge, replay-ability, ease of use, and entertainment among other factors. Additionally, the use of in-game quantitative data, specific to the use of BCI and other event related information has been collected for analysis.

# 3 Experiment Design

## 3.1 Experiment Overview

The overall research question of whether the Microsoft Kinect and the Emotiv EPOC worked together as a combination of input technologies was the primary goal for this project. The other goals determined to be influential in how this project was structured included:

- determine if the inclusion of BCI with the Microsoft Kinect provides any modification to the: entertainment, challenge, replay-ability, and other similar attributes.
- determine if using the BCI is a complete hindrance to gameplay or if it co-exists naturally.
- determine which brain waves captured by the Emotiv EPOC brain device will be useful for determining the qualities required to play the outlined games.
- determine how readily individuals can control their own brain waves for their own success.
- determine if there are any unrecognised implications in the development of games that involve the Microsoft Kinect and the Emotiv EPOC BCI device.

To accomplish the evaluation of these goals three different puzzles were created. Each puzzle was designed to test different uses of the inputs. It was necessary to keep in mind some important factors when designing the gameplay experience. The participants would for the majority not have had any exposure to any form of BCI equipment. The previous experience with voice and gesture recognition systems like the Microsoft Kinect could not be guaranteed either. Therefore, as the participants would only be involved for a single session the gameplay had to be simplified to maximise the playability and decrease the learning curve with the devices.

The application provides a set of three different puzzles. These puzzles have been named as follows:

- Tile puzzle
- Street puzzle
- River puzzle

When each participant was ready they would start the game from the main menu and complete each of the puzzles listed above in the order provided. Each puzzle is designed to test the participant in a different way and force them to “think” differently. While completing puzzles, actions were recorded for later analysis to provide a more diverse set of results. Post analysis of data collected could then be conducted after the completion of all puzzles. For all the puzzles the participant was be provided with an information screen prior to the gameplay (Appendix C: Back Story contains examples of these information screens). This screen set the scene providing the goal to the user and explaining what they must do to complete the puzzle. These screens are brief to not overwhelm the participants. The instructions for each puzzle were primarily explained verbally and participants were given the option to read the information screens as well if they chose to. Once the participant confirmed they were ready, the relevant puzzle would begin.

At any point during the game, play could be paused by the user with the voice command "main menu". This provides a popup that provides the functionality to exit to the main menu. This feature was not indicated to participants as the play through should be contiguous. It was required though so that the researcher observing the session could exit out of puzzles where the participants were unable to complete a particular challenge.

The puzzles have been loosely joined together with a logical storyline. The tile puzzle sets the scene and has the participant trying to unlock a secret chamber of a tomb by solving a map. Once they have gained entry they discover the tomb has already been raided, but spot someone escaping. This is where the player must chase the thief in the streets but not alert them. The goal is to find out where the treasure has been hidden. However, when the hideout has been found the thieves throw the treasure into the river while trying to escape. During the river puzzle the task is to then collect as much of that treasure as possible so that it won't be lost.

As part of this experiment overview a discussion of a number of different elements that come together to form the overall experiment are considered. Firstly the choice of hardware is identified discussing the use of the Microsoft Kinect and the Emotiv EPOC hardware with why they were selected. This is followed by an overview of how the experiment was set up when testing with participants, and the procedural structure of the tests. Then an overview of the design of the individual tasks the users had to complete including a calibration stage, and each of the three puzzles. Finally, an overview of the questionnaire and how that was designed to fit in with the tests is summarised.

### 3.1.1 Choice of Hardware

The Microsoft Kinect and Emotiv EPOC devices provide only two instances of specialised equipment. There are a variety of other technology solutions available that accomplish individually similar feedback. This includes solutions for detecting movement input using the Playstation Move of Wii Remote for example. Or in the case of a device that can be used for reading brain inputs the Neurosky (Neurosky, 2012) is another example. The Microsoft Kinect has been popularised by its connection to Microsoft's Xbox360. It is only recently that support has been provided for the Kinect to be used for Windows based games and other development. There is a version of the hardware that is separate for each environment (PC and Xbox360). The Xbox based version was used for this project as that was the hardware that was available. The differences between hardware for the Kinect are not significant for this project. The reason the Microsoft Kinect was selected for this project is primarily due to a number of factors:

- The Microsoft Kinect includes skeleton tracking, allowing for tracking of relative position of hands and other body parts for use in puzzle solving.
- The Microsoft Kinect includes voice recognition, allowing for speech control for menu navigation and a method for solving puzzles.
- The integration with the XNA based game development environment is simplistic as both are developed by Microsoft, and this was always to be the preferred development environment for the project.

In contrast the Emotiv EPOC device was chosen as an EEG device based on availability of the hardware within the University. The Emotiv EPOC hardware has 16 electrode nodes for reading EEG information and 2 additional inputs that can measure a two-axis gyro for determining the orientation of the headset. The Emotiv EPOC headset's name suggests the reading of an emotional state. The software that comes with the headset includes measuring of a number of different emotional states. Some of the states include: Engagement/Bordeom, Frustration, Mediation, among others. (Emotiv, 2012) For the purposes of this project it was decided that the numbers calculated by the Emotiv EPOC software related to the specific emotions would not be used. This was primarily due to the proprietary formulas used to calculate the emotional states were not provided. The alternate deployed algorithm will be described later in section 3.2.4. The method that was used in this project essentially used a rescaled mean of a pair of EEG node inputs. The NeuroSky is another headset available for public purchase, but unlike the Emotiv EPOC with its variety of sensors, the Neurosky only provides a single electrode as an input. (Neurosky, 2012) Based on the complexity of the brain a larger number of sensors provide a larger set of data that could provide additional insight to findings and future research. Therefore the choice of the Emotiv EPOC over the NeuroSky as the hardware of choice for this project made sense.

### 3.1.2 Experiment Setup

Participants were taken through four different stages of the experiment:

- The first stage ensured participants understood the requirements and were given the opportunity to ask questions.
- The second stage calibrated the software to the individual. By performing a simple task, a scale was generated to be used in ensuring the software responded in a similar way to different users.
- The third stage involved the three different puzzles. The puzzles were completed in the same order each time. The Tile puzzle provided a task that had no time limit, thus ensuring a chance to become used to using the input technology. The Street and River puzzles require more active movement with a time constraint.
- The fourth and final stage asked users to complete a questionnaire to provide feedback about their experience.

The first stage does not have any additional design context to further explain the way the experiment was carried out. For each of the other stages the following sections will describe the design of the individual components.

Participants for the experiment were found by approaching students and academics at Flinders University. It was initially anticipated that a group of 20 to 30 participants would take part in the experiment. Many of those who had initially indicated they were available and willing to participate were unable to attend sessions for a variety of reasons. This meant that only 15 participants were included in the experiment. The results therefore should be considered as part of a preliminary study.

The hardware used in the experiment was the same set used for all tests. This included two networked laptops via an Ethernet cable, the Emotiv EPOC, and the Microsoft Kinect. As seen in Figure 5 below, the participant would be asked to keep a distance from the Microsoft Kinect sensor and the display, while wearing the Emotiv EPOC.

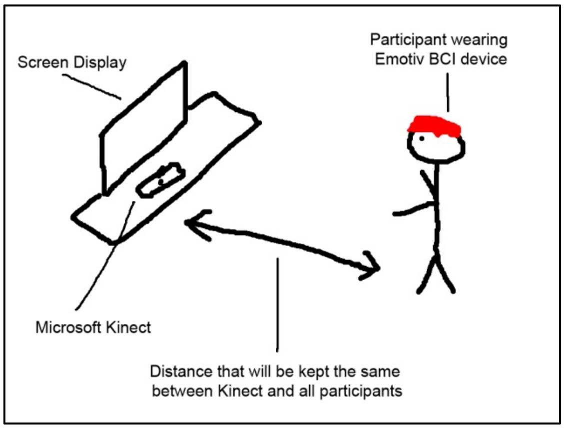



Figure 5. Test Setup

The reason for using two networked laptops is that one had the license for the Emotiv EPOC software, and a license for using Matlab. The laptop with these licenses had insufficient hardware to manage running the game, so the data collection from the Emotiv EPOC device was separated from the game. As indicated earlier, a Matlab based application was developed by Dr Sean Fitzgibbon, another researcher at the University, to facilitate reading input directly from the device and sending it via the network to the application designed as part of this research. The application also stored the full collection of data during each session storing it locally as an additional form of data storage.

### 3.1.3 Puzzle Design: Calibration Stage

The specifics of the software implementation of the calibration will be discussed in section 3.2.4. The calibration stage was introduced to give the participants a visual method of calibration. In the early stages of development the calibration stage was completed entirely on the server side. This meant the early development builds required a messy setup, with limited feedback. This was less than ideal. Especially as it was impossible using this method to determine if the calibration was successful until already entering into the active game. If the calibration had failed then the game had to be terminated and the calibration restarted. It was particularly for this reason the calibration stage was transferred to an in game solution.

The primary goal of the calibration stage was to accomplish two fundamental elements for immersion:

- Firstly and most importantly to improve the gameplay of the later puzzles by simplifying or increasing the difficulty. Some participants depending on a large number of factors will exhibit varying levels of achievable relaxation or concentration. By performing a calibration it will provide some element of standardisation.
- Secondly to introduce the participant to the basics of the Emotiv EPOC BCI device.

The calibration functionality introduced the participant to the idea that they would have to relax and then concentrate to perform different actions in the game. In Figure 6 below, the calibration window is shown. To ensure the participant could not skip the steps all the interaction after entering this screen was managed by key presses rather than gesture or voice recognition. The initial view indicates a wait for instructions command for the participant. At this point it would be explained that they would need to first relax for a period of 20 seconds. It was explained that simply trying to not focus on anything or think about anything while attempting to relax gave the best results.

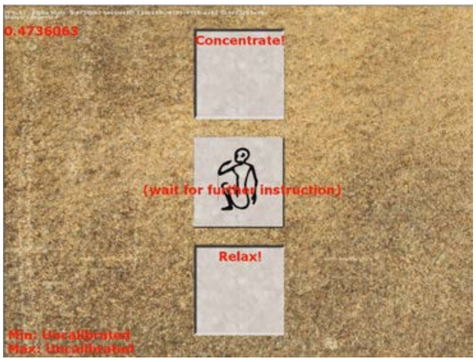



Figure 6. Initial Calibration Screen

Once the participant confirmed they were ready a key press would begin the calibration for relaxing. Relaxing gives values that are in a high range for Alpha waves. Since the calibration has not been performed a pseudo method of indicating that the participant was relaxing was devised. The tile that can be seen in the middle of the screen would move down toward the "Relax!" box as the user is appearing to relax (the text "Concentrate!" would not be visible during this stage). This can be seen in Figure 7. The large numbers in the top left corner of the window show the raw Emotiv EPOC data as it is fed through. Subsequently higher values result in a downward translation for the tile. So if the values were 0.18 followed by 0.25 then the tile would move down, and would continue to move down if followed for example by 0.38. When a number that is smaller than the previous value occurs, for example a 0.23 following the 0.38, the movement of the tile is halted. This feedback is expected to demonstrate to the participant that they have not maintained the necessary state. Once the 20 second timer finishes the "Max" value shown in the bottom left will change from "Uncalibrated" to show the upper scaling value. This is calculated by using the values received over the entire period of the 20 seconds. The formula used will be discussed in section 3.2.4. The value is found by taking the upper quartile of the set of 20 values taken during the 20 seconds.


Figure 7. Performing the Relaxation Stage of Calibration

The next stage won't begin until a key press is performed. During this intermediate period the max value can be checked and determined if the value is high enough. In the case that it is too low or too high a reset button is available to clear the calibration back to the initial stage. It is not expected that every participant is able to immediately command their mental state effectively so this is an important inclusion. Before beginning the concentration calibration stage the participant can be instructed that they should attempt to concentrate and focus on moving the tile up. This will again, when completed, provide positive feedback indicating to the participant they are performing the correct action. An example of the concentration stage much like the relaxation stage can be seen below in Figure 8.

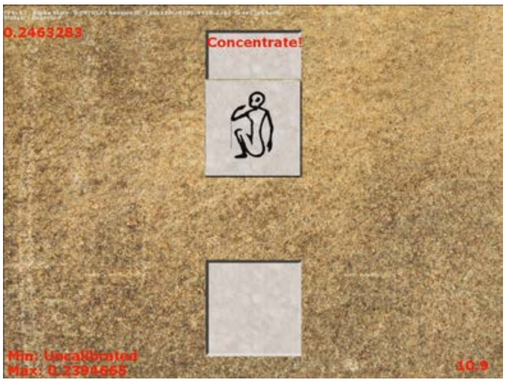



**Figure 8. Performing the Concentration Stage of Calibration**

After completing the 20 second period of concentrating a "Min" value will appear in place of the "Uncalibrated" wording. The functions related to this will be discussed later, but the function uses the lower quartile of the 20 data elements to select a value. At this point it can be determined if the min and max values are suitable, specifically that they have a reasonable range between them. Numbers that are too close typically indicate a failed calibration. Again, this is why there is an included reset mechanism to allow recalibration if necessary. After the min and max have been set the number shown in the top left will no longer be a raw number from the processed Emotiv EPOC data, but a rescaled value. The final stage of the calibration allows for movement of the tile in both directions based on both the concentration and relaxation. This is only a necessary stage if the min and max values are considered to be potentially unsuitable. It allows for a few seconds to check that the values being presented represent useful data. In most cases when testing this was simply skipped over as it was not necessary.

### 3.1.4 Puzzle Design: Tile Puzzle

The aim of the puzzle is to arrange a series of tiles in the correct position and orientation. The participant is presented with a three by three grid of tiles initially showing little or no means of determining how they are related to each other. The aim of the puzzle is to arrange the tiles in the correct locations within the grid with the correct rotation. To accomplish this, the user should relax and the real, “hidden” images should become gradually more visible. The images that appear will indicate a correct method of ordering the tiles by location and rotation. Once the tiles have all been correctly aligned the user will be presented with notification that they have successfully completed the puzzle. In this particular puzzle a map of the world was chosen as the hidden image. It could have been anything that was distinct in showing a correct orientation and position for each of the individual pieces. Figure 9 below shows the world map used. In particular the compass point was added in the top left and the border around the outside to reinforce the correct positioning. The divided map can be seen in Figure 10 below, which includes the backdrop of the tile graphic used to decrease the visibility of the original tile while still retaining the original appearance.

Figure 9. World Map

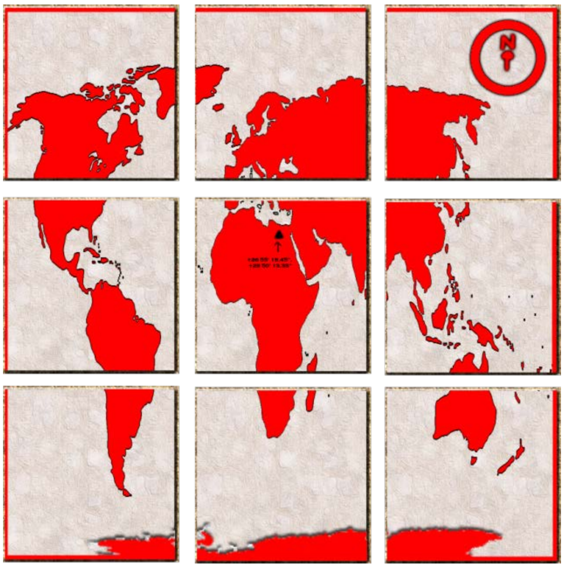


Figure 10. World Map Divided as Tiles

This puzzle relies heavily on physical, verbal, and brain inputs. The physical interaction involves movement of the left or right hand to perform the actions of a cursor for selection of onscreen elements. The verbal input provides cues for interpreting the commands: “swap hands”, “drop tile”, “select tile”, “rotate tile right”, “rotate tile left”. There are a couple of variations for each of the first three commands to accommodate a more natural vocal input for the user. These commands will allow interaction independently from the reliance on the brain input. For the participant to make sense of what they are doing logically the brain input provides this extension.

The voice recognition provided by the Microsoft Kinect is definitely not perfect. It was for this reason and allowing participants to also have a choice, that the variations of verbal inputs were included. Table 1 lists the phrases that a participant can use to perform the various actions listed. The imperfections of the voice recognition will be categorised further in the results; however, it is worth indicating a couple of examples of incorrect actions. For some users when clearly pronouncing “Drop Tile” the speech recognition software interprets the commands to be “Grab Tile”. And another example is the word “Tile” in the middle of “Rotate Tile Left” and “Rotate Tile Right” can often be left out without any negative effects with the terms correctly interpreted. Therefore the errors can provide both advantages and disadvantages for the participant potentially.

| Operation | Verbal Phrase | Alternate Verbal Phrase |
| --- | --- | --- |
| Select the tile that is currently underneath the hand cursor and attach it to the movement of the hand. | Select Tile | Grab Tile |
| Drop the tile that is currently held into an empty cell or swap the tile with another tile below the currently held one. | Drop Tile | Place Tile |
| Toggle the hand currently being used to control the position of the cursor. (Will toggle between the left and right hand) | Change Hands | Swap Hands |
| Rotate the currently selected tile left or right. (The two listed are not alternates they perform the specified verbal action). | Rotate Tile Left | Rotate Tile Right |

**Table 1. Tile Puzzle Phrases**

As stated earlier, the brain input uses a method to determine the relaxation level of the individual playing. By continuing to relax, the overlay images will become increasingly visible with a changing opacity. It was originally thought that concentration would be ideal for being the determining factor to reveal the hidden images. After initial evaluations early in the development of the application it was decided that relaxation would be a better indicator. The reasoning behind this decision is that while solving the puzzle the participant will need to be concentrating anyway. Therefore, logically there would be limited potential for fluctuation in the BCI data to provide useful gameplay experience. Requiring the participant relax to make the image appear is believed to require a change in state or the way they think. As they relax the image will appear. Then when the participant needs to shuffle the tiles around and decide the correct position and orientation their concentration will cause the image to fade away again. To indicate the gradually increasing and decreasing of relaxation, an animated bar is used as seen in Figure 11 and Figure 12 on the following page. Figure 11 shows the initial state of the tiles with no visible information about the correct order and orientation. As the participant reaches a continued relaxed state over a period of time the bar will have filled up to the state shown in Figure 12. In this figure the image is shown fully revealed. It was decided that the tile underneath should still be visible to provide increased visual recognition for when the hidden world map image disappears again.


Figure 11. Tile Puzzle Initial State


Figure 12. Tile Puzzle Revealed State

In terms of the anticipated neural activities the participant will have a number of influences skewing their mental state. They will need to be concentrating on how their chosen left or right hand is orientated on the screen and how it relates to the cursor, while also trying to get the speech recognition software to understand them. The failure of speech recognition, in understanding a participant's commands, will likely cause frustration. Frustration will likely decrease a participant's ability to relax. The frustration would therefore increase the difficulty in making the hidden map appear. Through all of these actions the participant will need to attempt to relax. It is entirely possible that participants can maintain a revealed map the entire time. For those who struggle, it is possible based on the way the puzzle is designed to relax until the image is visible and then begin moving the tiles around while the hidden image fades away. Then repeat the relaxation step again as required. The drawn out nature of the puzzle is the reason why it was chosen as the first puzzle for participants to be tested with. It can be thought of as a training period before they are expected to have to interact faster for subsequent puzzles.

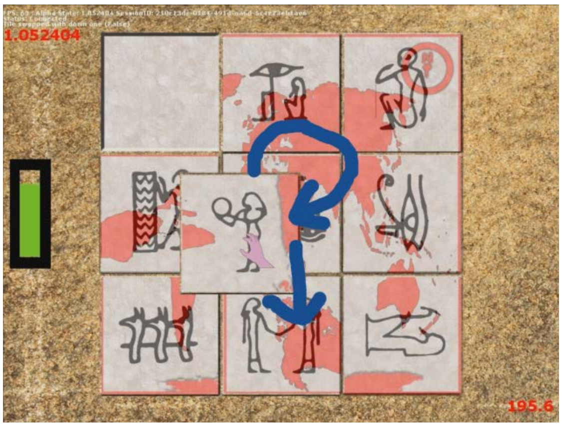



Figure 13. Transition requiring a "rotate tile right" and a "drop tile"


Figure 14. Tile Puzzle Complete

To conclude the design overview of the Tile Puzzle there are two additional figures that need to be addressed. Above in Figure 13 is shown an example of a tile attached to the cursor and the steps for the tile to be rotated and then placed. This also shows how the region underneath the tiles appears. After successfully completing the puzzle, the success is indicated to the participant by the dialog seen in Figure 14 above. The dialog allows for returning to either the main menu or proceeding to the next puzzle.

### 3.1.5 Puzzle Design: Street Puzzle

This puzzle places the user in a bird's eye view of an Arabian style marketplace. The aim is to chase a thief who is on their way back to their hideout where they have hidden the stolen treasure. This is to be considered a continuation of the story provided in the first puzzle. The participant must follow the thief back to the hideout without alerting them to their presence by getting too close. The player and objects are placed on a series of "rails" and the individual will move between these rails to avoid objects. Colliding with objects will slow the player down and risk getting too far behind. The thief will randomly stop from time to time and look around so the player must be weary of this and be ready to stop themselves and hide. The player's in-game character speed is conversely related to their relaxation state, for example the higher the relaxation the slower the speed. An example screenshot of the game is seen below in Figure 15.

Figure 15. Street Puzzle

Obstacles within each level are randomly placed based on a simple set of rules. A map is randomly generated with a path for the thief to travel and then object markers are randomly placed in the empty cells. When there are a whole row of cells with only a single cell break for the thief's path there is a chance of a river with a bridge segment being placed. Otherwise there is a random chance of market stalls, boxes, or other animated characters appearing. All obstacles give a negative 20% speed modification. This is enough to cause the player to lag behind if they continue to collide with obstacles. Table 2 on the following page shows a list of the obstacles and the graphics that were created to represent them.

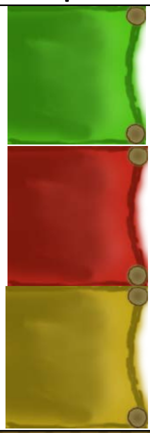
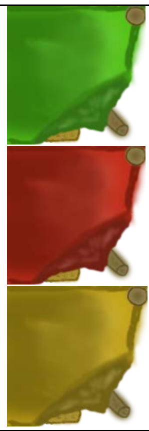
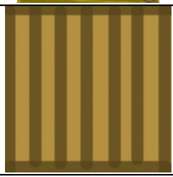
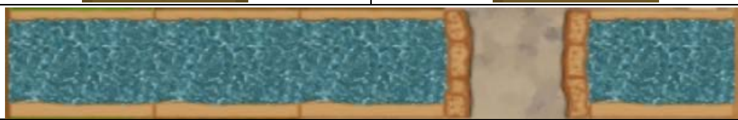
| Obstacle Type | Normal Representation | After Collision |
|---|---|---|
| Market Stall: This was made in three variations to provide variety. After colliding the graphic changes to what is intended to be the partially collapsed stall with the pole that was holding the tent part up fallen over. |  |  |
| Crate: Initially a solid crate, but once collided with a cracked hole is shown. |  |  |
| River and Bridge: The river and bridge combination doesn't have a collided variation. |  | |
| Bystander: The bystander will appear as a character with different coloured clothes to either of the thief and player and a random facing direction. All have speech bubbles that randomly appear over their heads to simulate the appearance of talking. After collided with their chat changes from "…" to "!!!" as if to appear annoyed. |  |  |

Table 2. Street Puzzle Obstacle Types

The input mechanics are simpler in this puzzle than the tile puzzle, but there is a higher element of risk that results from being detected by the target being followed; thus failing the puzzle. The physical input for this puzzle is primarily controlled using the centre of the hips as a translation point for determining which "rail" the player's character will appear on. As there is an included possibility of failing and the speed is modulated by the brain input a stopping mechanism was also included. Raising both hands above shoulder height will signal a stop of the player's movement. At times when the thief suddenly stops and waits for a couple of seconds this will be the only way for a player to stop themselves if they are already too close. Originally a failure mechanic was included for when a participant becomes too far away from the thief too, but to account for cases where a participant is unable to make their character speed up it was removed.

Unlike the tile puzzle the anticipated neural activities require what could be considered a lower amount of thought. The participant's focus is most likely to be on attempting to avoid objects by physically moving from side to side, and then realising the times when they have to slow themselves down and stop. For those who are travelling slowly it is possible they may become frustrated. Individuals who were able to grasp how to control the input reasonably well during the first puzzle are expected to succeed in this puzzle with ease due to the lowered amount to consider.

### 3.1.6 Puzzle Design: River Puzzle

For this puzzle the story is that the thief has thrown the treasure into the river in a bid to escape. The player's goal is to grab as much of the treasure that is floating down the river within a period of time. The problem for the participant is that there is a lot of junk floating down the river as well. As the player relaxes they will slow down time so that the items float past slower. The amount of time remains the same in a relative sense for the puzzle. The total time is 150 seconds. So if the participant managed to keep the river slowed to the slowest speed (0.5 times normal river speed), then it would take 300 seconds for the puzzle to end. And during that time the same relative number of items would be spawned. An example screenshot of the River Puzzle is seen below in Figure 16.


Figure 16. River Puzzle

The primary input uses the skeletal tracking of the Kinect to move an on screen cursor "hand" that will be used to collect the items that are flowing from right to left. To select items the cursor must be held over the item that is to be selected for 0.4 seconds as the object is moving. Then the items must be dragged into the bottom of the screen to deposit them into an "inventory". By concentrating and becoming relaxed the participant can optimise their collection activities by having the time to collect increasingly large amounts during the same period of time.

The physical input in this task should not be difficult unless there is some reason for a person playing to not be capable of raising their arm for a few minutes. It is anticipated that the primary cause for concern on the part of the participant is frustration from missing the collection of treasure. The frustration of missing objects as they pass by could negatively impact a participant's ability to clearly focus on maintaining the slow speed. A negative state will give increasingly more difficult gameplay. There is a maximum speed of 1.3 times normal speed to prevent the difficulty becoming too high.

### 3.1.7 Puzzle Data Collection

The information collected for evaluation and analysis can be broken into two categories. The first is for information that is common between all the puzzles and is therefore generic information. The second is the information that specifically relates to each of the puzzles. The general information that was collected for the first category included the following:

- X and Y locations of user hands/hips.
- Speech commands issued.
- The alpha wave and other wave states as a numerical list over time.
- The time taken to complete each puzzle.
- A unique identifier created as a GUID to represent each participant. This could be used as a reference for the questionnaire, server side logging, and in application logging to provide a link between the data.

For the second category each puzzle has relevant data that needs to be logged. Firstly the tile puzzle logged the following information for later review:

- Events that occur including:
    - The start of the puzzle,
    - when a participant performed (or attempted to perform) an action such as selection or dropping of a tile.
- The configuration of the randomised puzzle positions and a calculated minimum number of swaps required to correctly place and orientate the tiles for use in comparing whether the participant had performed acceptably.

Secondly for the street puzzle the following information was logged:

- The map and random path of the thief.
- Events for when a participant changes 'lane', stops themselves, and when the thief stops.
- The distance kept between the thief and player over time.
- The speed and effect influenced by the BCI inclusion.

Lastly for the river puzzle the following information was logged:

- The events relating to the successful collecting of items and junk compared with the player missing items. This would include the influence of the BCI as a numerical value to identify how easy it was to choose the item.
- The speed over time for how fast the items were flowing in the river based on the BCI input.

### 3.1.8 Questionnaire Overview

The questionnaire may be found in Appendix A: Questionnaire. It was designed with four sections:

- Establish participant background.
- Determine subjective evaluation of tasks using a Likert scale with rankings of 1 to 10.
- Provide written feedback regarding experiences.
- Complete a NASA Task Load Index.

Each of these was designed to provide useful feedback on different aspects of the research.

## 3.2 Software Design

### 3.2.1 Overview of Development Stages and Design Changes

The development of the software followed an agile methodology. Initially as a pilot development test, a Simon Says type game was developed to test the capabilities of the Kinect for speech recognition and skeletal input. This trial development stage resulted in the creation of almost the entire code required to manage the Kinect based input between all the different puzzles. After this it took 29 cycles of development to reach the complete and polished application for testing. The elements that still had to be added to the game were broken down into stages and expanded upon. The cycles were based on periods of time and so, design elements were shifted to the next cycle when they were not completed within time.

### 3.2.2 Software Design Overview

The overall design of the product developed for this project was that a participant could be viewing one window at a time. Windows may be considered the construct that is used to derive menus, puzzles, and other high level structures. All of these windows should have access to updates from all the input sources that they require, and be capable of indicating to the application when the window is finished with so that a change to the next view can occur. Some of the specific class based details will be discussed in the next section in regard to the class structure used in the product. The most essential aspect of the entire application based on the HCI component of this research is the way in which input has been handled. Everything except for a single special case for input in this product has been funnelled into a single input management class. The application is built to support keyboard input, Xbox360 gamepad input, Microsoft Kinect motion input, mouse input, and BCI input.

The special case that was mentioned is the speech recognition. It would have been possible to store the current state of the voice into the input manager. This was opted to not be completed this way as the event based nature that is highly targeted for a specific purpose made it difficult. All of the other inputs provide a numerical detail that could be needed by multiple different regions of code. To handle speech recognition instead it is handed to the current window if the current window has speech support to do whatever is necessary with the interpreted speech.

Keyboard, mouse, and gamepad inputs are included as methods of debugging. Except that the calibration stage locks input to use only keyboard. This was done to restrict accidental input by participants so that the researcher must handle the state changes. This will be discussed more in relation to the integration of the BCI in section 3.2.4. The following sections will continue to discuss the integration and use of inputs and controls for the application in regard to their design and implementation.

### 3.2.3 Class structure

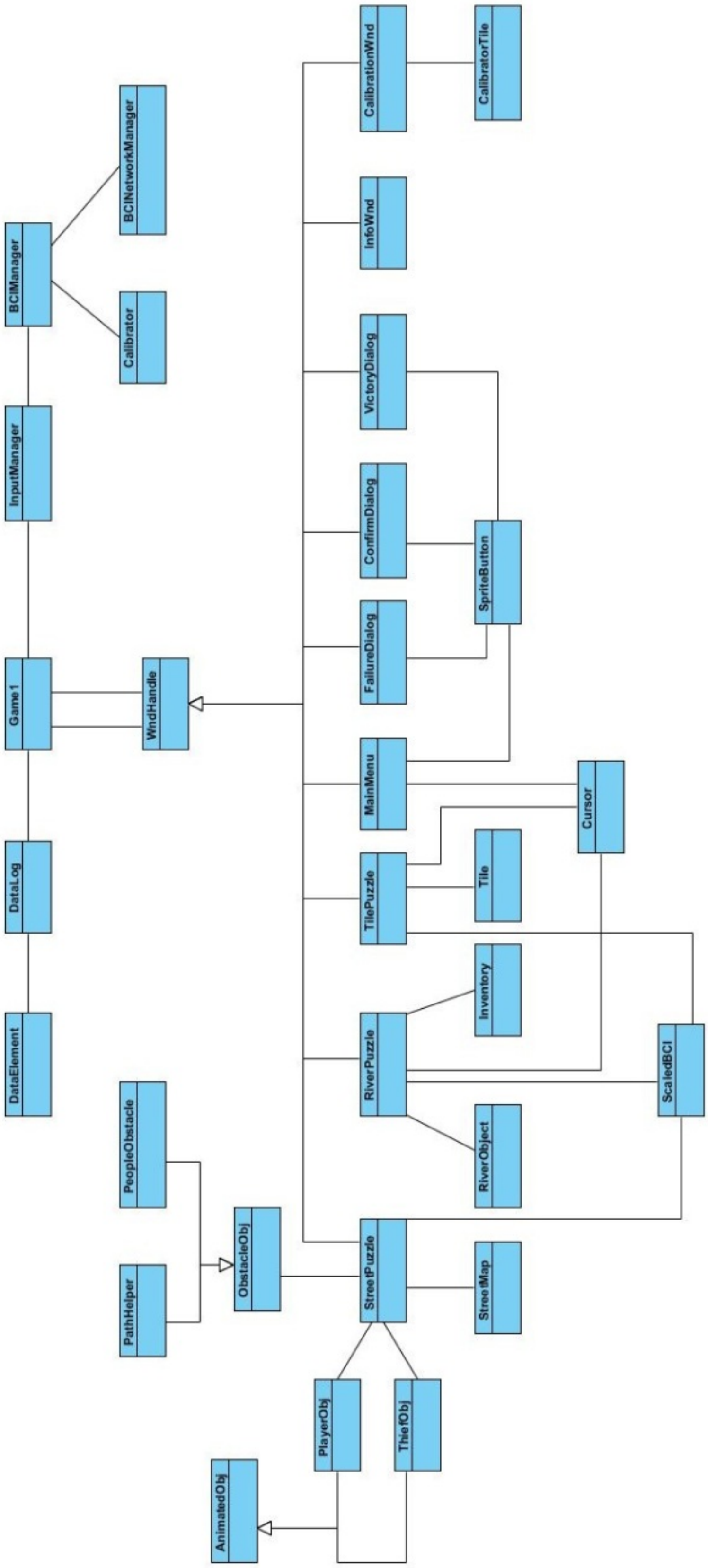



Figure 17. Class Diagram

The basic structure of the designed application is shown in Figure 17. The primary game class is simply named “Game1”. The WndHandle class is inherited by 9 different sub classes to provide the window support that is used by “Game1”. These related classes include:

- “RiverPuzzle”, “TilePuzzle”, and “StreetPuzzle” are all puzzles.
- “MainMenu”, “InfoWnd”, and “CalibrationWnd” are all other supporting windows.
- “FailureDialog”, “VictoryDialog”, and “ConfirmDialog” are all windows that are designed to be used as overlays on top of the puzzles.

The other significant details in relation to the classes are that the windows all have access back to the primary application through their WndHandle. Using this reference they gain access to the “DataLog” that provides the logging mechanism used to store all the results from this research, and InputManager that provides the access to all the current input information.

### 3.2.4 Integration of BCI

Originally the BCI component of this project was to be a simple augmentation to be combined with the inputs from the Microsoft Kinect. Access to the Emotiv EPOC was only obtained from around halfway through the project. The game was developed with classes to handle the appropriate BCI interaction, but allowed for simulated values to be provided through keyboard or Xbox360 controller input. The use of these additional debugging inputs was invaluable throughout the project. Configuration of the Emotiv EPOC takes time and based on experience while testing for extended periods of time it becomes tiring to use after a while. This could be due to the strain of having to think about being in a state of relaxation or concentration. This particular aspect will be discussed later in the results discussion section of this document.

The first stage of working with the physical device was to work toward having the live data being streamed from the device to the game. The Emotiv EPOC software is licensed to Flinders University and the necessary software installed on a specific laptop. For this reason the laptop had to be used as a separate processing tool. The laptop’s hardware was insufficient to run the game so a pair of laptops was required. This ruled out direct integration of the code to access the device into the game code. Instead a client-server configuration was designed. The laptop reading the data from the Emotiv EPOC takes the data and stores it locally; it then passes a frame of data over the network. Through consultation with Dr Fitzgibbon this server application for the Emotiv EPOC device was developed using MATLAB.

### *3.2.4.1 Initial BCI Implementation*

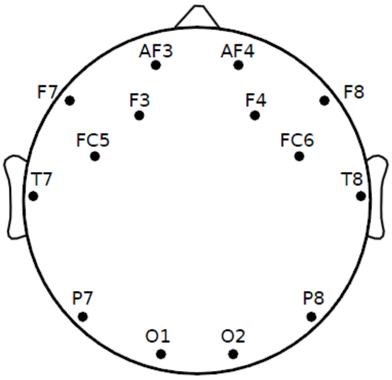



**Figure 18. Positions of Electrodes for the Emotiv EPOC (Martinovic, 2012)**

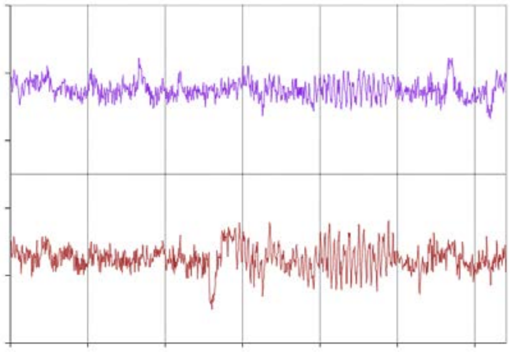


**Figure 19. Sample EEG showing O1 and O2**

Initially the application was built to follow the data flow diagram as seen above in Figure 20. The MATLAB based server had the capability of recording any number of calibrations. Each calibration would be additive to the previous calibrations. The function of the calibrations would take the minimum and maximum values read and use those as a rescaling upon the data when sending it from the server. Calibrations would take a period of 20 seconds. To perform the calibration the individual using it would have to press a calibrate button and then relax (often completed by closing of the eyes during these early stages) or concentrate. There was no visual feedback indicating the usefulness of a calibration. This then became a significant issue as the calibration had to be configured prior to activating the server, then if an issue arose and it was necessary to recalibrate, the server had to be closed down. As the calibrations took the absolute maximum and minimum values this did cause issues that made a poor calibration appear more dramatic when it occurred. The data that was sent across the network by this application included the entire frame of 8 bands and for each band the 16 channels resulting in a total of 128 values. All of these values were the raw readings from the device without any scaling applied. The two values that were placed at the front of all these were the calibration scaled numbers considered useful; specifically: posterior alpha using the channels O1 and O2 from the 10-12 Hz band and frontal theta using channels F3 and F4 from the 4-7 Hz band. The positions of these registers can be seen on Figure 18. A sample of EEG is seen in Figure 19 with mild concentration (low amplitude), and mild relaxation (high amplitude) in the second half.

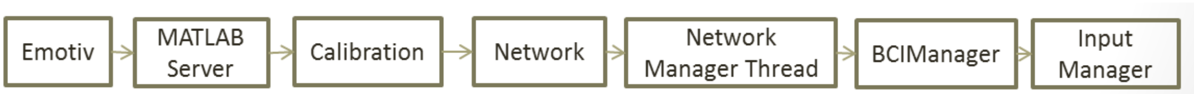



**Figure 20. Original BCI Data Flow**

The game that connects to the server as a client controlled this interaction by running a separate network management thread as identified in Figure 20. The job of this network thread is to continually listen to incoming data from the server. It stores the data in variables that can be accessed through an interface by the BCIManager class. Its other purpose is to handle when there is no connection to a server. When this happens, the BCIManager class is notified of the issue and the network thread will continue to reattempt connection to the server periodically. Once successful connection is made again the BCIManager is told that the server is live once again and operation can resume as normal. The BCIManager's task is to handle all of the values that have been passed across the network external to the thread used by the main game loop. The BCIManager is referenced by the InputManager. The InputManager allows the interfacing for all the inputs to the application including the BCIManager to allow simple control. Using this initial scheme there was no transformation of the values on the client side. The values would be directly applied to each of the puzzles. This introduced a number of issues including dramatic changes from one value to the next, and the threshold was inconsistently useful between calibrations reducing the overall effectiveness.

#### 3.2.4.2 Intermediate Considered Strategies

A variety of strategies were considered to counter the issues that arose as part of the initial development. The most prominent solutions were developing a dynamic threshold system and using a neural network to match against patterns representing the states. The use of a neural network was immediately ignored based on the additional time required. At the time of development it was drawing close to when participants would need to be found and tested. For this reason dynamic thresholds were explored instead.

The first designed implementation of this was based on the principle of high values being desired (high values equate to a relaxed state). When an individual reached a peak value above a threshold the value obtained would be retained for a period of time (5 seconds was used as a test case). Additionally the issue of dramatic changes between values was solved by using a smooth transition between updates in the standard increment of 1 second of time. When a value higher than the current state was hit this algorithm would boost the retained state to that level and then hold that for a new period of time. A dynamic threshold was trialled along with this that reduced the threshold every 30 seconds to simplify the gameplay. The primary issue with this approach was that it did not allow for simple handling of both concentration and relaxation based gameplay. Also, the gameplay became too easy or conversely, with minor changes to the thresholds, it would suddenly switch and become impossibly difficult. Finding the correct balance using this approach was not working. So a final series of changes saw the final model that was developed.

#### 3.2.4.3 The Final BCI Model


Figure 21. Final BCI Data Flow

The final model's data flow can be seen in Figure 21. It differs from the original model by the placing of the calibration stage after the data has been collected by the BCIManager and the inclusion of a ScaledBCI class. The calibration stage was moved from being a server side operation to solve the complexities with having to imagine relaxation and concentration. Instead a visualisation could be provided and rapid checking for the correctness of a calibration could be completed too. To calculate the calibration the series of 20 values over a period of 20 seconds would be collected. These 20 values were then sorted and the upper quartile for the relaxation based maximum value and the lower quartile for the concentration based minimum value were taken. This was selected as a method to complete the calibration as it would ideally cull out any extreme values. The maximum and minimum values were used to rescale the numbers that were sent across the network before passing them to the other classes that required the input. The Calibrator class was used to apply a calibration to the values when requested from the BCIManager. To minimise the influence of large muscle activity, values over 3 were culled and the previous lower value used in place. This number was determined based on the values that appeared with eyes closed, so that it was just above this. This meant that an eyes closed method could still be used potentially by participants or to allow for the normal muscle contamination that might occur. This summarised method for calibration is shown in Figure 22.

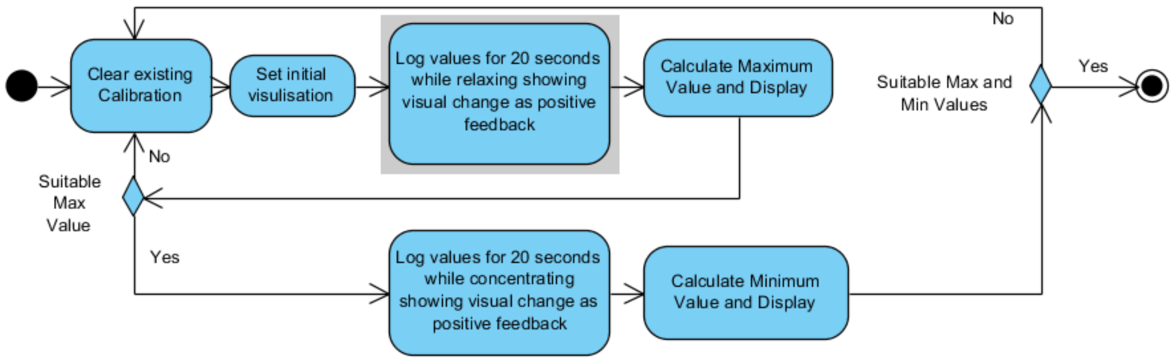



Figure 22. Calibration Activity Diagram

The ScaledBCI class is so called as it provides a scale between a minimum and a maximum value. The configuration of these values as well as a threshold for concentration and relaxation can be specified on a per puzzle basis. The passing above or below the related threshold enables an increment or decrement that is applied to the scale moving it toward the minimum or maximum values. This general application flow can be seen in Figure 23 and would be executed inside the update method. The values that are passed to this class are assumed to have been passed through the Calibrator class to use the scaling determined in the calibration stage. It is significant that the values represented by this class incremented over a period of time instead of instantaneous jumps between intervals. The period of a second was used to match the rate at which data was being sent from the server. This enabled a transition period of a second to represent a smooth effect. This is a particularly simplistic model that could be adapted for use in potentially any game. The model as seen in the diagram has been made more complex by allowing definitions that can provide benefit from either concentration (low values) or relaxation (high values) by using negative numbers. The threshold used for the puzzles was 0.4 with the expectation that values higher than this would correspond to relaxation and values lower to concentration. The threshold for each could be modified independently, and there was an extension to the class designed that allowed for dynamic thresholds to be defined. This was excluded after initial trialling of the code as it did not appear to be particularly effective for any real purpose. This is likely because the method used was to change the threshold based on time independently of how well a participant was performing. An approach that could be used to further extend the ideas presented here is to process the participant's active effectiveness and determine how well they are performing to estimate whether a threshold change should be applied. A similar approach could also be applied to the calibration to actively enhance the output of using the device.

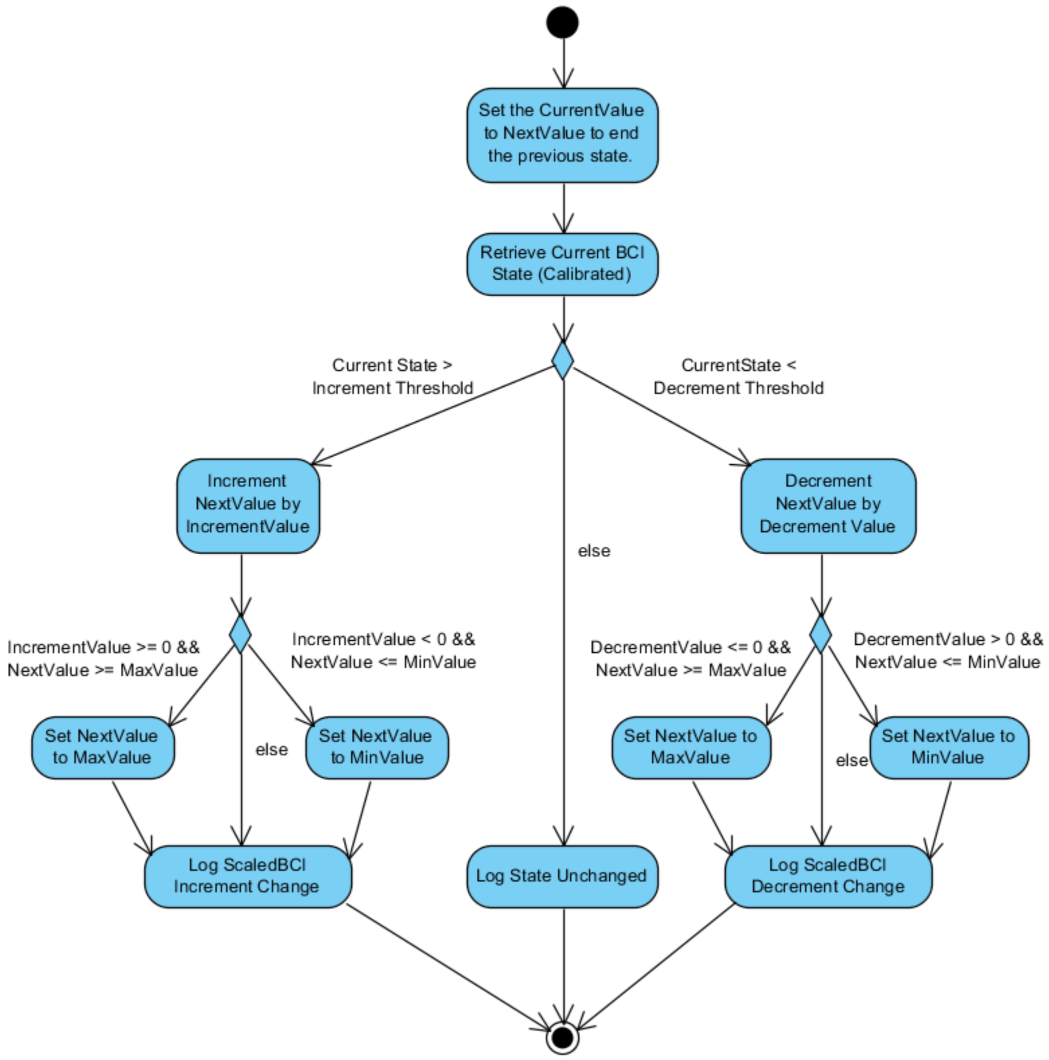



**Figure 23. ScaledBCI Activity Diagram for Update Event**

### 3.2.5 Integration of Microsoft Kinect

Within the puzzle games a number of different Kinect based interactions had to be addressed and tracked appropriately:

- One or both hands.
- Centre of the hips.
- Centre of shoulders.
- Verbal commands to provide event driven interaction.

In the Tile Puzzle the player could use one of the hands and choose to interchange at will between them using a verbal command (“change hands” or “swap hands”). The position of the user’s hand is tracked by the Kinect and then converted into a screen coordinate to control the movement of an in-game cursor. The tiles could be selected using “select tile” or “drop tile”, rotated using “rotate tile left” or “rotate tile right”, or dropped using “drop tile” or “place tile. Much of the interaction for this was developed by creating a “Simon Says” style game with colours that could be selected. The “Simon Says” game utilised the same input mechanics described above: the user’s hand location was tracked and translated into a position to represent the cursor on screen, then verbal commands could be issued to perform the selection of coloured tiles in the correct pattern. The mechanics of this type of puzzle were evolved into a puzzle that could more appropriately be augmented by the use of BCI.

The Street Puzzle required tracking of left and right hands, centre of shoulders, and the centre of hips. The hands and shoulder combination were used as a stopping mechanic, such that when the hands were both higher than the shoulder centre the player’s character would stop. The centre of the hips was used to determine the lane the player’s character was walking down. The River Puzzle only required a single hand input, but could be changed just as it could in the Tile Puzzle with “change hands” or “swap hands”. This was used to hover over moving items to select them and transition those items into the inventory.

### *3.2.5.1 Handling of Speech recognition*

All the terms that can be recognised by the speech recognition software API (Microsoft Speech API was used) have to be registered and parsed prior to gameplay that needs them. For this reason each window that required commands was designed to hold a collection of these terms. When the speech recognition system is being configured along with the Kinect all of these terms are pooled together with duplicates removed. Using this automated process made it easy to modify and change the terms that could be said. The speech recognition API allows applications to handle the events for: speech recognized, speech hypothesized, and speech recognition rejected. Speech recognized is triggered when speech has been successfully hypothesized. Until then, the API indicates its current state by passing messages via the speech hypothesized event. And if the hypothesized word could not be matched to a phrase the speech recognition rejected event occurs. In this application the recognised and rejected cases were both recorded; however, only the recognised event had full support for feedback to the user. The WndHandle class provided an interface for handleSpeechRecognised(string s) and handleSpeechNotRecognised(). Once the string had been passed through, the string could be matched to the correct action if it was relevant to the current context. Similarly if the term was not in the correct context (for example saying “select tile” when inside the Street Puzzle), the term would be automatically ignored as an incorrect phrase. The general flow of how speech is handled can be seen in Figure 24.

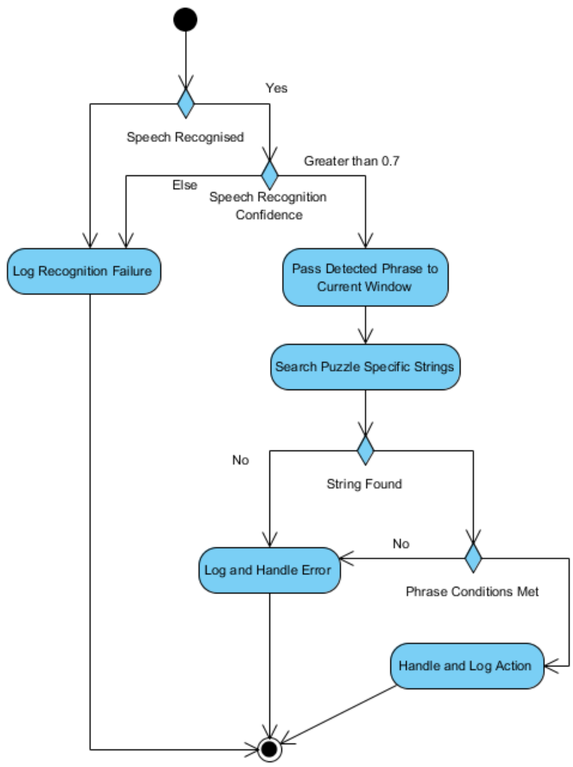



**Figure 24. Speech Recognition Activity Diagram**

#### 3.2.5.2 Handling of Skeletal tracking

Skeletal tracking with the Kinect is event based. To simplify the potential complexity of having the specific event having to post a message to all windows, the information required was stored in the InputManager. This allowed a transition from event based model to the update and draw model used by the rest of XNA. The information being tracked was determined based on the settings in the InputManager. For example if only one hand was needed then only the position of the specific hand was calculated and stored. These requested modes could be enabled and disabled by the specific puzzles. Importantly all the coordinates received by the application from the Kinect API were transformed into coordinates matching the screen space. This allowed for simple direct mapping to positions within the screen. However, it was found during development that for hand movement this was not far enough. The hand movements needed to be exaggerated with a large physical movement to complete actions, so a second rescaling is applied in puzzles that need the hand inputs to reduce the physical movement required. This particular aspect is discussed further in the results section of this document.

#### 3.2.5.3 Speedier debugging methods

Using the Microsoft Kinect for continuous active debugging is difficult in a tight space; or even more so when the device itself was not available. For this reason the debugging methods were designed to allow continuous development of the application unhindered by any requirement to use a real Kinect. To accomplish this, the Kinect's existence was checked and a flag used in the InputManager to indicate the availability. Additionally an InputMode could be specified if one specific input type was to be used. Then a debug input method could be run in any of the puzzles or other windows to simulate input. For example a key press could be used to signify saying a phrase. As an approach to development with this form of HCI equipment it was superior for rapid redeployment.

### 3.2.6 Other Design Considerations

The primary other design consideration that has not already been discussed is that the main menu includes a rendered depth frame. The depth frame that the Kinect provides as an event when it is ready allows determination of the acceptable distance and angle the player needs to be from the Kinect sensor. It uses the infrared camera to determine the distance of objects in the scene and highlights objects using this distance. As this was included in the main menu it allowed rapid checking that the distance participants were standing from the device was satisfactory.

# 4 Results

## 4.1 Results Overview

The testing phase of this research had 15 participants. Ideally there will be additional participants in future studies. This initial group have provided a useful set of preliminary results. As seen in Figure 25 below 47% of the participants were in the age range between 21 to 30 years of age. The participants who were under 21 were also over the age of 18. Only 2 of the participants were female, so there was not a fair distribution of both genders. The largest issue with female participants, even for those who participated is the amount of hair many of them have. To correctly position the Emotiv EPOC on the head it requires as close to skin contact as possible. Large quantities of hair complicate this.

80% of the participants were students with the remainder indicating they were either academics at the university or other staff members. 87% of participants indicated they spend more than 31 hours of time on a computer each week as seen in Figure 26. This indicates the majority are heavy computer users.

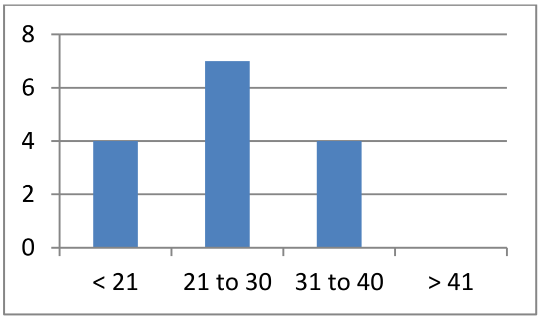



Figure 25. Participant Ages

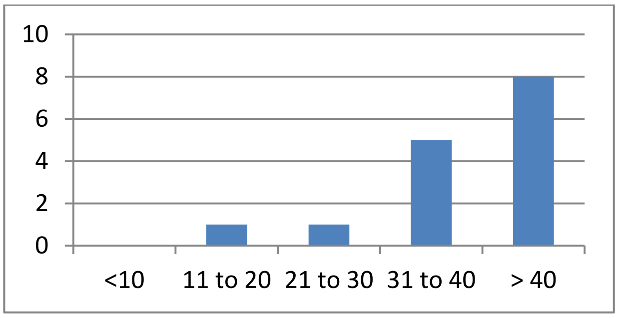



Figure 26. Time spent on Computer

In Figure 27 below it is shown that 67% of participants indicated they spent less than 10 hours a week playing games. In many of these cases it was also indicated that the respondent's time was taken up by University and they have been more active gamers in the past. In Figure 28 the range of motion technologies participants had indicated exposure to previously are shown, 73% indicated they had used a Wii remote, while 60% indicated they had previously used a Microsoft Kinect. The use of motion technologies can be compared against the use of consoles. 93% had used a Playstation previously, but only 33% had used the Playstation Move (shown as PSMove on Figure 28). 67% of participants had used Xboxs, with 60% having used the Kinect. All except one participant who had used an Xbox had previously used the Kinect as well. Finally only two of the participants had previously had experience with BCI equipment. Both of these participants were academics.

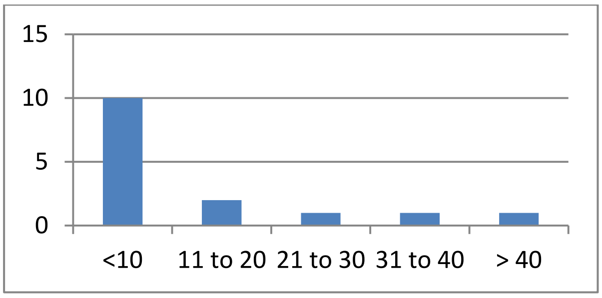



Figure 27. Time spent Gaming

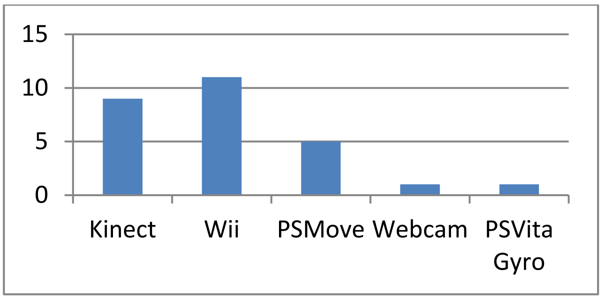



Figure 28. Use of Motion Technologies

For a slight majority of participants they passed through the puzzles without any difficulties. 40% of participants were unable to complete the Tile Puzzle. For 5 of these it was due to failure by the speech recognition API. The speech recognition API refused to correctly interpret anything these individuals were saying. To ensure this was not due to some other cause the experiment observer stepped in to test a couple of verbal commands. The phrases spoken by the researcher in all except one of these cases were correctly interpreted. Significantly, both female participants were in this group of unrecognised individuals. The remaining participant who was unable to complete the challenge was finding it too difficult to make the hidden image appear. Additional information that has been extracted from the collected results of the application will be reviewed in section 4.3.

## 4.2 Questionnaire Results

The information about the participant backgrounds obtained from the questionnaire has already been reviewed in the overview of this section. The remaining sections of the questionnaire that participants completed were the Likert 1 to 10 scale questions, written responses to clarify their thoughts, and the NASA Task Load Index.

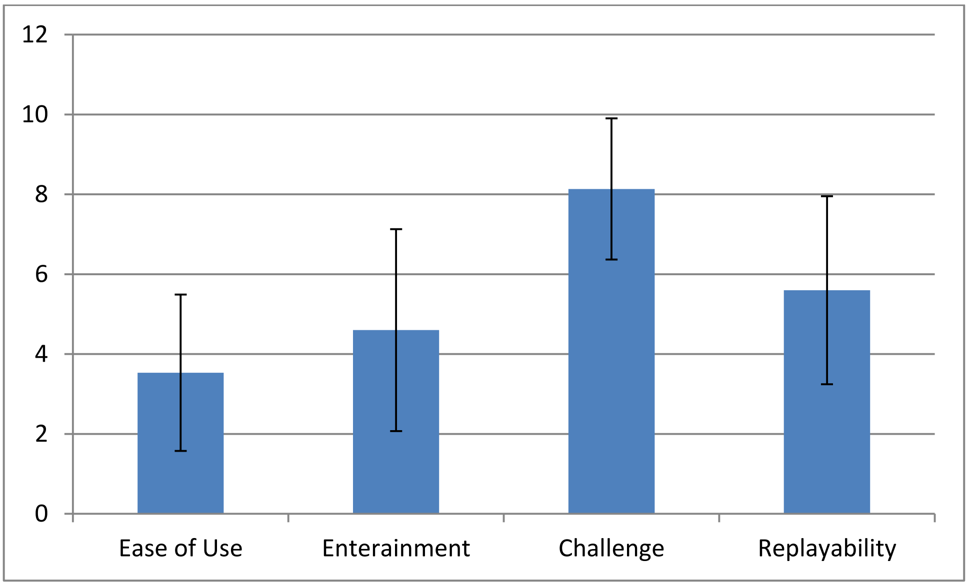



**Figure 29. General Gameplay Responses**

Figure 29 above shows that participants indicated on average there was significant challenge found in the puzzles. The ease of use was marked as being particularly low, and a medium level of entertainment. The replay-ability was responded to with a larger score though. This suggests that although there was an inherent challenge and difficulty using the system, the participants had indicated an interest to play the puzzles again.

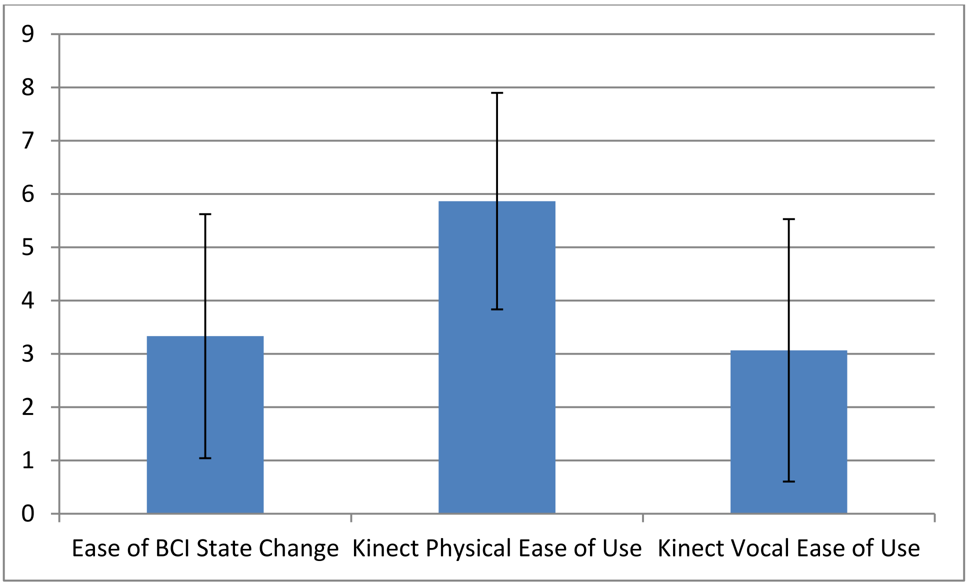



**Figure 30. Ease of use with specific HCI Components**

In Figure 30 the responses to the ease of use with the individual HCI components is shown. Participants found it particularly difficult to change their state between a relaxed and a concentrating state. In comparison, the physical input for the Microsoft Kinect was rated far higher. Interestingly the speech recognition was rated lower than the BCI; however, this would have been heavily influenced by 6 participants finding it impossible to complete the Tile Puzzle challenge.

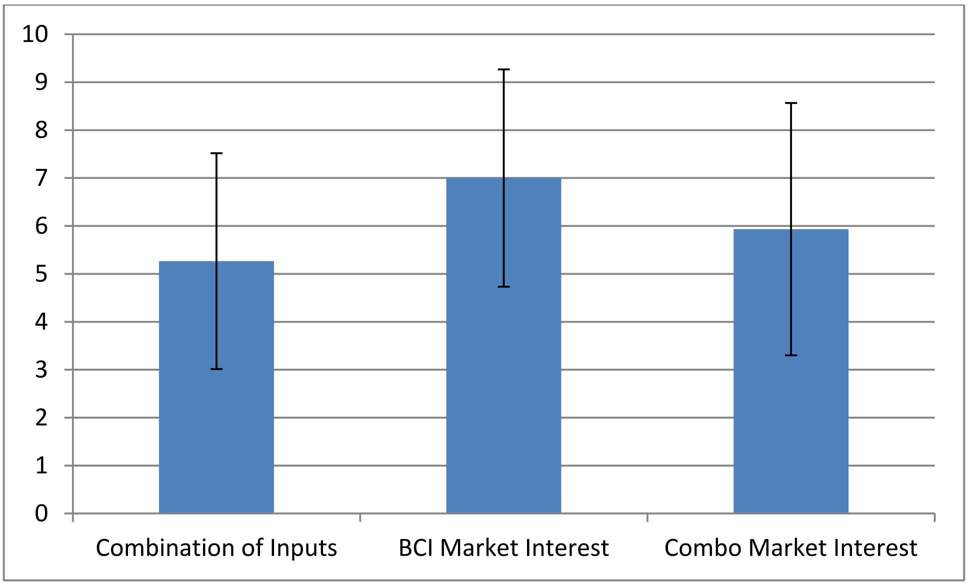



**Figure 31. Combination of Inputs and Marketability**

In Figure 31 the combination of inputs was rated only around 5.3 on average. Suggesting that there is a lot of work yet to be done if the combination of BCI with gestural and speech inputs are to be marketed. There was a high interest in BCI being brought into the market by individuals. The Emotiv EPOC (Emotiv, 2012) and NeuroSky (Neurosky, 2012) are both BCI systems currently available to consumers. The hardware has not been incorporated into any major gaming releases as a primary input source at the time of writing though. There was interest shown in the combination of inputs being provided in a commercial setting. This was contradicted by a slightly lower interest in the general combination of inputs. This seems to be an irregularity, as it seems illogical the interest in buying the combination is higher than the interest in the hardware being combined.

The results from the NASA Task Load Index ranged between 61 and 80.7 with a mean average of 69.1. This does indicate a reasonably high load for the task as a whole. The participants indicated the highest areas of difficulty were the mental challenge and frustration.

## 4.3 Application Collected Results

Between the raw brain data stored by the server and the data collected by the application there was over 1.1GB of data collected relating to the events that occurred within the game. For this reason only some of the key results will be reviewed that stood out. The data recorded was enough, such that if desired a replay could be run by feeding the recorded data into the application. At this stage the application has not been built to have this functionality, although it is certainly a feature that could be added later if further testing were to be completed with it. An automatic data mining application was constructed to extract the data used for the results analysis.

### 4.3.1 Tile Puzzle

It has previously been said that 40% of participants were incapable of completing the Tile Puzzle. 5 of those 6 were unable to complete the puzzle due to inability for the speech API to recognise their voice. Then for the other individual it was due to an inability to maintain a state of relaxation. Due to the heavily skewed results depending on the success or failure a few of the metrics taken from this puzzle have been separated based on the success. As seen in Figure 32 the participants took an average of 983 seconds to successfully complete the puzzle. For those who were unable to complete the puzzle an average of 593 seconds was spent attempting to make the puzzle work prior to switching to the next puzzle.

In Figure 33 the average voice recognition failures is shown. In cases where the participants were successful there was still an average of 4.2 failed recognitions detected. This was far higher for the participants who were unable to have the application register their voice at all at an average of 15.8 failed recognitions. There was only one case where the participant completed the puzzle with no failed recognitions. In this same case the participant also completed the puzzle in the minimum number of moves possible. The highest number of failed recognitions was 29.

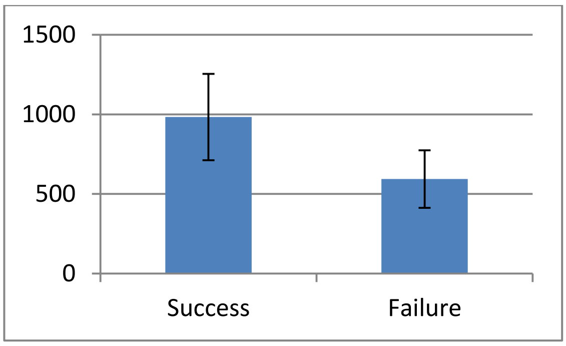



Figure 32. Average Puzzle Completion Time for Tile Puzzle

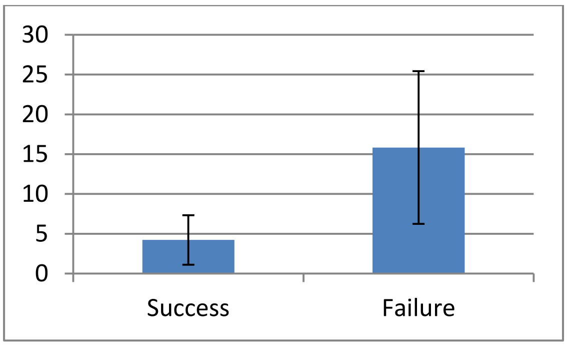



Figure 33. Average Failed Voice Recognitions

As seen in Figure 34 the average time spent concentrating was 29% (rounded up from 28.75%) and the time spent relaxing was 71% (rounded down from 71.25%). The aim of the puzzle required entering a relaxed state so this does indicate that on average the puzzle was either providing too little challenge or that there were larger issues that must be considered. It is suspected that the primary reason is due to muscle movement contaminating the data. This will be discussed further in section 5.2. This hypothesis is supported by 40.4% of values being repeated. A repeated value for the application can occur in three cases:

- The very rare occasion when the BCI state is exactly the same twice in a row.
- The application has automatically ignored an input as it is above the maximum threshold. (The maximum threshold was set based on a standard eyes closed state). Typically caused by muscle movement such as clenching of the jaw.
- The BCI has not successfully posted information across the network within the time frame and the repeated value is due to lack of updated data.

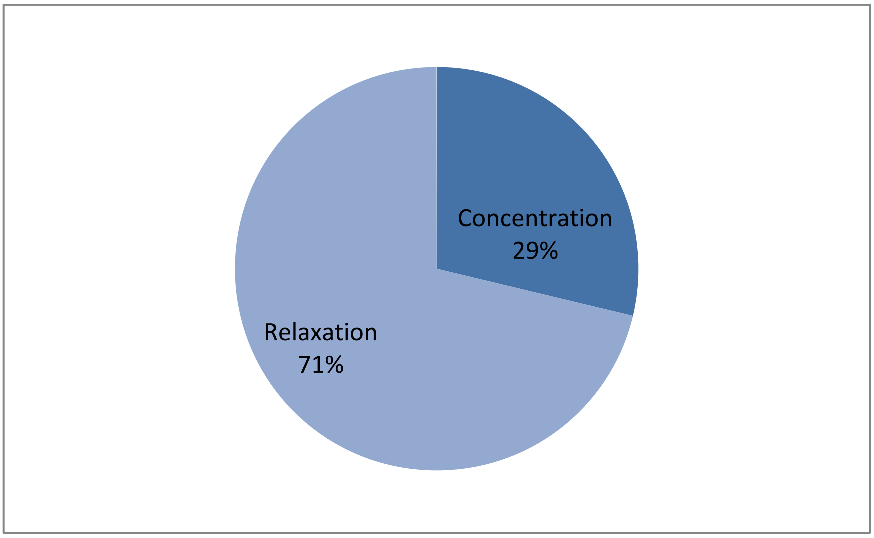



**Figure 34. Average BCI States for Tile Puzzle**

For the individuals who were successful in completing the puzzle the average number of moves that was calculated to be the minimum were 27.8 moves. Each selection, drop, and rotation is considered one move. The average number of moves taken though was 38.4, resulting in an average of 10.6 additional moves required. This high difference can be attributed to a couple of different factors. The first factor is there were a couple of participants who found the voice commands did not work in all cases. "Rotate tile right" would fail as a voice command leaving "Rotate tile left" as the only choice. It was not a programmed feature, but many participants left out the word "tile" from the phrases and they were still correctly identified. Only three participants chose to swap the hand used for control in the puzzle at any point. The other factor was the proportion of time the puzzle was visible. Due to the inability to see the puzzle at times, participants made guesses that were incorrect for the placement of tiles.

### 4.3.2 Street Puzzle

The Street Puzzle like the Tile Puzzle did have a small failure rate. For the Street Puzzle it had an automatic reattempting system built into it to allow for up to three retries. Three participants failed the puzzle on their first attempt due to staying too close to the thief for too long. These three passed the puzzle on their second attempt, so every participant completed the puzzle. As seen in Figure 35 the ratio of relaxation to concentration was exactly the same for this puzzle on average as it was for the first puzzle. Concentration accounted for 29.2% of the time, and 70.8 % of the time for relaxation. The repeat rate was slightly higher for this puzzle at 41.8% of values being repeated. The highly movement orientated nature, having the user move from side to side, may be to blame, the other reasons would be the same as suggested for the previous puzzle for example, that it may either be the rare case of the same input appearing, or the network delay causing a delay for the updating throughout the rest of the application. On average the puzzle took 351 seconds to complete. With times ranging from as high as 484 seconds to as low as 297 seconds. On average 20.5 objects were collided with on the 100 rows of objects that could be collided with. During gameplay an average of 452.5 lane changes were recorded for participants. This particularly high number is likely due to an error in the game that causes positions the player can stand in resulting in rapid toggling between lanes.

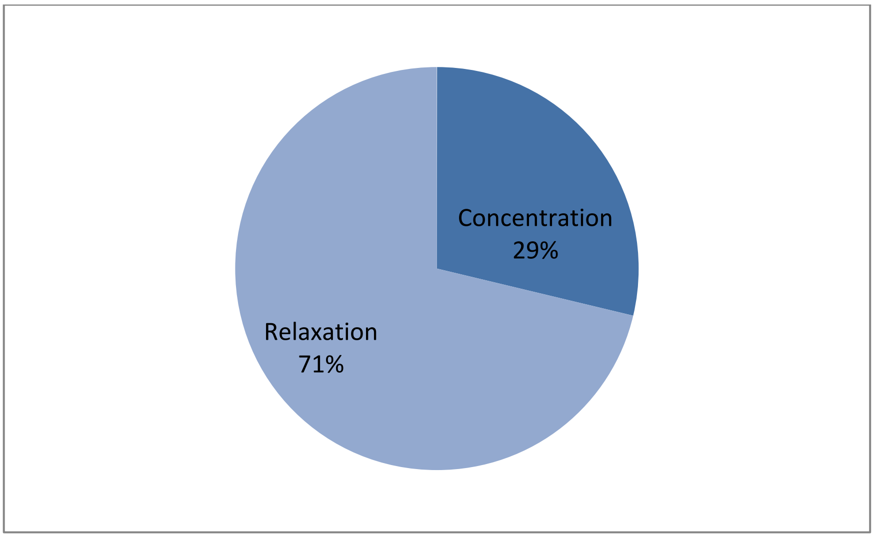



Figure 35. Average BCI States for Street Puzzle

### 4.3.3 River Puzzle

The River Puzzle was a challenge that was impossible "fail" in any sense, but provided a potentially difficult scenario to accomplish with a good score. The results for the distribution for concentration and relaxation are particularly skewed for this puzzle. In terms of concentration 67% of participants had 13.4% concentration time or below. Of those participants 4 had 0 concentration readings. The two highest readings of 87.6% and 100% are definitely outliers in this regard; however, it is important to recognise that for all three puzzles the concentration values were rated at a particularly high level for these participants. The combined average for this puzzle gave 22.7% of the time spent in concentration and 77.3% of the time spent in relaxation as seen in Figure 36. The high muscle movement may be to blame for the high values. Only 39.3% of the values were repeated for this puzzle which was slightly less on average than either of the previous puzzles. On average participants took 167.8 seconds to complete the puzzle. This indicates on average most were viewing the puzzle in normal time as the puzzle was designed to take 150 seconds and then slow or speed up time based on the BCI modifier. Only 6 participants accidentally collected up to 2 pieces of junk. Therefore the target acquisition rate was high, it was explained to participants that the yellow objects were the ones they wanted so this is not a surprising statistic. On average 19.5 treasure objects were collected with a maximum score of 49 items. The three high scores of 43, 47, and 49 corresponded to the three longest times. The remainder ranged between 5 to 26 collected items, indicating a lot of frustration was had by some; particularly based on the number of objects missed. An average of 38.9 treasure objects were missed with a minimum of 17 and a maximum of 58.

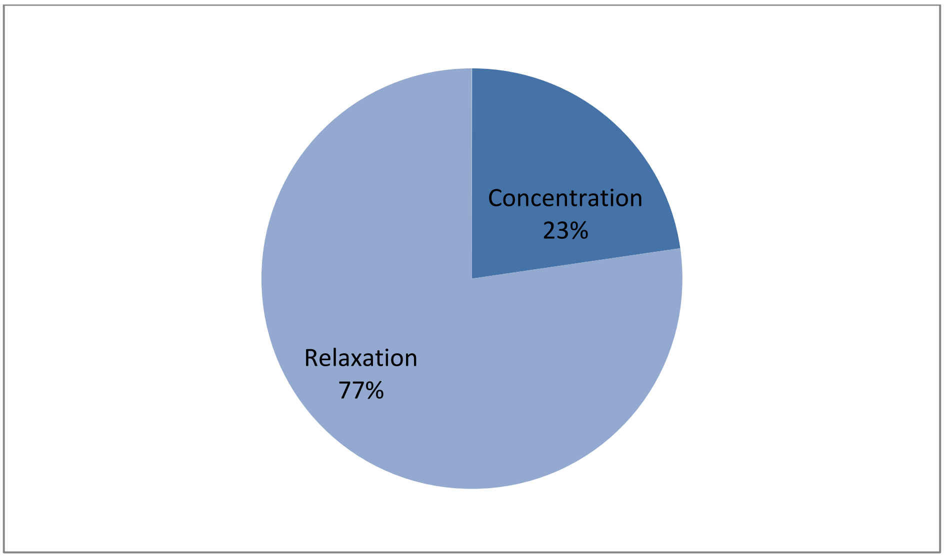



**Figure 36. Average BCI States for River Puzzle**

# 5 Results Analysis

## 5.1 General Observations

The high average for time spent in a relaxed state undermines the purpose of the puzzles. As may be seen in Figure 34, Figure 35, and Figure 36 there was a very similar high percentage of time spent in a relaxed state. Figure 37 shows this as a summary comparing the results of all three for the proportion of time spent concentrating, relaxed, and for the time repeated. The intent for the first puzzle was that a relaxed state must be obtained. Many maintained this “relaxed” state once it was achieved despite the obvious appearance of concentration in determining the puzzles. From observation there were a number of cases where the puzzle worked exactly as intended. Of more concern than the results from the BCI state analysis is the number of participants for whom the speech API failed to recognise their speech. In addition to the information logged there was observed often a build-up in errors. The application for the first puzzle has a verbal feedback message spoken to the participant indicating the application did not understand what was said. At times the application let out a build-up of errors and a continuous stream of verbal messages saying that the application did not understand what the user said. The entire processing side of the application that handles those messages is controlled from within Microsoft’s API. The triggering of the message would occur directly after the speech rejected event was fired.

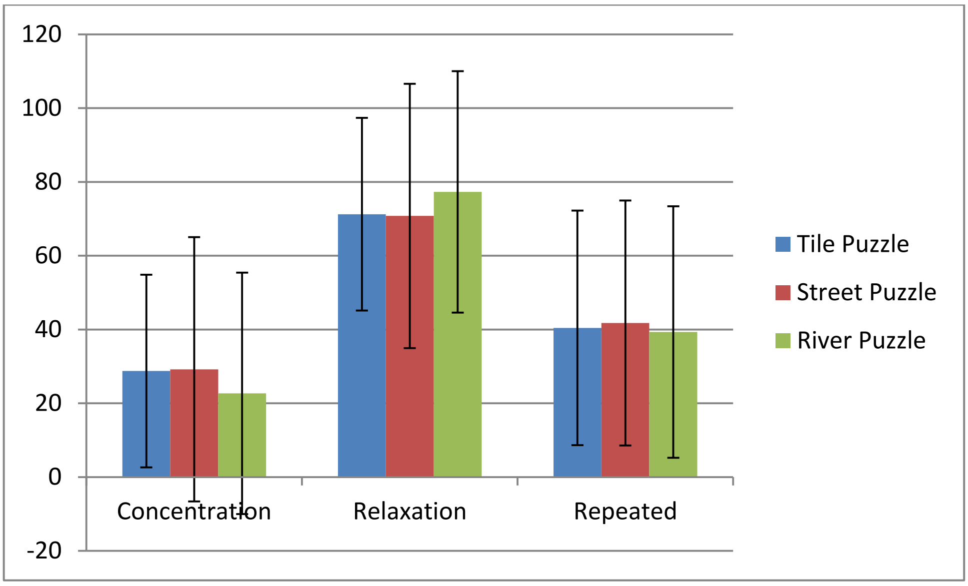



Figure 37. Comparison between BCI for all three puzzles

## 5.2 Issues in Analysis

### 5.2.1 Microsoft Kinect Technical Issues

In the case of the Microsoft Kinect there were only a small number of technical issues cited by observation and individuals. As the hardware is marketed with large games like Mass Effect 3 (EA International, 2012) it would be expected that the device should be robust and highly capable. During the middle of this research a new revision of the software was released by Microsoft that claimed to fix many issues, so it is possible that some of these issues have been mitigated.

- Speech recognition sketchy and slow to respond to phrases at time.
- Skeletal input was jittery at times. (A stationary hand would appear to jump around the screen)

To mitigate the issues that were discovered a number of strategies had been implemented. Unified input management by combining inputs into a singular management object allowed simplified extraction of information when it was required. Multiple phrases were used for significant puzzle goal completion. For example “grab tile” and “select tile” as two dramatically different sounding phrases were implemented to enhance the chance that at least one phrase would be detected correctly. This was seen as a required element for some participants to pass the challenges as “rotate tile right” did not work for one individual. And for the researcher “drop tile” was often misinterpreted as “grab tile”. Finally to minimise the complexity and strain of hand movement the hands positions were rescaled to increase the sensitivity.

### 5.2.2 BCI Technical Issues

The technical issues encountered with the BCI equipment certainly slowed development of the application. The project was intended to use the BCI as a simple augmentation using existing research to identify required inputs to then apply the output to gameplay. Some of the issues that were encountered with the device include the following. Examples of these issues are:

- Difficult to know the battery state for the wireless headset as the metre shown in the provided software did not work.
- Required correct placement on the head with ample contact and a reasonable level of moistness that at times required a lot of manoeuvring or reapplication of saline solution to achieve.
- The headset appeared difficult to reliably control although after a lot of use it appeared to become easier.
- The calibration stage can make or break the game depending on whether participants provided reliable input that mapped correctly to their mental state throughout testing.
- One second delay between updates from the BCI server may have a detrimental effect on the usefulness of the data.
- Muscle movement heavily contaminates the data. For example a jaw clench appeared to give values 100 times what would normally be expected.

During testing the best that could be done was to ensure the device was powered and to attempt to get the electrodes past hair when there was a lot. It is perhaps the case that for this first trial the time with the device was insufficient to properly become used to using it. The Tile Puzzle’s deliberately slow pace was designed to slightly counter this by providing a chance to understand how to control their EEG outputs. The calibration stage as it is in the final game attempts to resolve the issue through culling of high and low values and the extreme values are culled actively within gameplay. The one second delay was chosen to limit the amount of data being sent across the network; by either combining the software into a single package or by further testing with lower delays this aspect could be improved.

### 5.2.3 Neurological Issues

This preliminary experiment anticipated a low learning curve for using the BCI equipment and Microsoft Kinect input. A couple of participants stated they would have preferred to have had the opportunity to use the devices separately to become used to them first and then brought them together. As the period of time participants used the software was typically for no more than 20 minutes their exposure to the devices was quite low. It may be an area for future research to investigate the amount of time necessary to become comfortable with the basics of using the combination of inputs or even simply the BCI component with any other form of input.

The other neurological issue that must be considered is the assumption of the application that the calibration stage will bind a set of fixed values that will be used for the collective puzzles. In situations where the participant was not in as heightened state of relaxation as they might otherwise be easily able to achieve it would make the game too easy and likewise for the reverse making it too hard. To resolve this issue in future implementations, a dynamic threshold system would be implemented to rescale and provide active recalibration. This could be done periodically or as a continuous process.

# 6 Future Work

There are a number of areas of development that can be explored for future work for the applications of the research. The most directly related area is the continuation of the work presented in this project. This project has presented a possible solution to combine the Emotiv EPOC and Microsoft Kinect. There was only one session each participant had the opportunity to experience using this combination of devices. Therefore one area that could be explored further is trialling either the puzzles completed for this research or other types of puzzles with multiple sessions. This would allow determination of any learning curve associated with the use of the combination of inputs and more specifically for the controlling of the BCI. It is felt by the researcher that from extended use of the Emotiv EPOC a greater level of ability to control the device has been attained, but without further investigation this aspect can't be proven.

Another area of expansion for this project is to compare the input devices used and to trial different devices. For example the BCI component of input could be compared against the NeuroSky (Neurosky, 2012), another BCI device available on the market. The Wii remote or Playstation Move could be used to potentially replace elements of the Kinect's gestural input. Additional methods would need to be used to accommodate skeletal inputs other than hands and the speech recognition provided by the Microsoft Kinect. Another area for expansion for this application would be to provide a multiplayer experience. The difficulties experienced in the single player puzzles could make players feel less disadvantaged if they are competing against another individual. This approach for BCI can be seen already in products such as the Mindflex Duel (ThinkGeek, 2012).

Games are only one of many areas for potential future development using this combination of inputs. Erp et al. (2011) provided a list that was presented in the literature review of areas for where BCI might be included. This list included: control of devices, user state monitoring, evaluation, training and education, gaming and entertainment, cognitive improvement, and safety and security. Although the full body type experience provided by integrating BCI with the Kinect is a more specialised type of input it could be used for enhanced natural user interfaces. Consider the interfaces presented in movies such as Minority Report (Dreamworks and 20$^{th}$ Century Fox 2002). Minority Report is an example of a science fiction movie including both a multi-touch natural user interface and the use of biometrics as forms of computer input. With the combination of inputs demonstrated by this project the potential for these forms of science fiction interfaces may be realised. Returning to the gaming aspect it could also be developed further for working with people who have different forms of physical and mental impairment. Natural user interfaces allow for expression based input and could allow sufferers of debilitating illnesses to experience a different form of enjoyment.

# 7 Conclusions

The research goals of this project focussed on specific questions. It was discovered that for the particular set of challenges presented to participants the entertainment provided by the inclusion of the BCI with the Microsoft Kinect was only a medium level (a rating of 4.6 out of 10 on average). The issues that have been identified as part of the use of Microsoft Kinect and Emotiv EPOC including speech recognition difficulty, muscle data contamination, and overall reliability indicate that there is much future work to be completed in this area. This pilot study has shown there are potential scenarios these two HCI devices could be combined in the area of games and entertainment. Alternate improved puzzles that utilise multiplayer, more stimulating puzzles, and improvements to the processing of BCI inputs could extend the entertainment provided.

Though the puzzles presented simple gameplay elements, participants reported a low ease of use but a high level of replay-ability. This aspect would be significant for potential future experiments. The experiments could compare multiple sessions to allow participants to experience the use of BCI for longer and determine learning curve information along with using a wider range of games to determine the types of games best suited for this type of interaction technology.

This research presents a basis that others could use for future evaluation of the role of BCI and NUIs. Alternatively it could also be used as an example of how a framework can be constructed for developing future BCI games that involve additional inputs like the Microsoft Kinect. The difficulties that have been identified can be taken into consideration by developers of software targeted at games or any other area of research that this combination of inputs could be used in. Until consumer BCI equipment reaches a stage when issues like muscle data contamination are not an issue it is unlikely BCI can be anything more than an augmentation. It is up to developers to consider how best to use this in fun and unique ways that can inspire the future of these input technologies.

# 8 Bibliography


[1] Bos, D.P.-O., Reuderink, B., van de Laar, .B., Gürkök, H., Mühl, C., Poel, M., Heylen, D., Nijholt, A. 2010, *Human-Computer Interaction for BCI Games: Usability and User Experience*, 2010 International Conference on Cyberworlds.

[2] Bolt, R. A. and Herranz 1992, E. Two-handed gesture in multi-modal natural dialog. In *Proceedings of the Proceedings of the 5th annual ACM symposium on User interface software and technology (Monteray, California, United States, 1992).*

[3] Cutrell, E., and Tan, D. 2004, BCI for passive input in HCI, viewed 24 March 2012, <http://hmi.ewi.utwente.nl/chi2008/chi2008_files/cutrell.pdf>

[4] Dreamworks, 20$^{th}$ Century Fox 2002, [Film] Minority Report, USA.

[5] EA International 2012, [Video Game] Mass Effect 3.

[6] Ebrahimi, T., Vesin, J.-M., and Garcia, G. 2003, *Brain-Computer Interface in Multimedia Communication*, IEEE Signal Processing Magazine, January 2003.

[7] Emotiv 2012, Emotiv, viewed 20$^{th}$ October 2012, <http://emotiv.com>.

[8] Fitts, P. M. 1954, *The information capacity of the human motor system in controlling the amplitude of movement*, Journal of experimental psychology 47(6): 381-391.

[9] Gentile, A. Santangelo, A., Sorce, S., Vitabile, S. 2011, *Novel Human-to-Human Interactions from the Evolution of HCI*, 2011 International Conference on Complex, Intelligent, and Software Intensive Systems.

[10]Güriök, H., Nijholt, A., and Poel, M. 2012, *Brain-Computer Interface Games: Towards a Framework*, IFIP2012.

[11]Graimann, B., Allison, B., and Gr¨aser, A. 2007, *New Applications for Noninvasive Brain-Computer Interfaces and the Need for Engaging Training Environments*, BRAINPLAY 07 BCI and Games Workshop at ACE, 2007.

*[12]*Hauptmann, A. G. 1989, *Speech and gestures for graphic image manipulation*. *SIGCHI Bull., 20, SI 1989, 241-245.*

[13]IJsselsteijn, W., de Kort, Y., Poels, K., Jurgelionis, A., and Bellotti, F. 2007, Characterising and Measuring User Experiences in Digital Games, International Conference on Advances in Computer Entertainment Technology 2007.

[14]Jacob, R. J. K. 1996, Human-computer interaction: Input devices. *ACM Comput. Surv., 28, 1 (Mar 1996), 177-179.*

[15]Kabbash, P., Buxton, W. and Sellen, A. 1994, *Two-Handed Input in a Compound Task*, viewed 24 March 2012, <http://research.microsoft.com/en-us/um/people/asellen/publications/two%20handed%20input%2094.pdf>

[16]Lecuyer, A., Lotte, F., Reilly, R.B., Leeb, R., Hirose, M., and Slater, M. 2008, *Brain-computer interfaces, virtual reality, and videogames*, Computer, 41(10), 66–72, (October 2008).

[17]Martinovic, I., Davies, D., Frank, M., Perito, D., Ros, T., Song, D. 2012, *On the Feasibility of Side-Channel Attacks with Brain-Computer Interfaces*, USENIX Security ’12.

[18]Molina, G.G., Tsoneva, T., Nijholt, A. 2009, *Emotional Brain-Computer Interfaces*, IEEE 2009.

[19]Niedermeyer, E., Da Silva, F. 2004a, *EEG-Based Brain-Computer Interfaces*, Electroencephalography 5$^{th}$ ed., Lippincott Williams & Wilkins.

[20]Niedermeyer, E., Da Silva, F. 2004b, *Neurocognitive Functions and the EEG*, Electroencephalography 5$^{th}$ ed., Lippincott Williams & Wilkins.

[21]Niedermeyer, E., Da Silva, F. 2004c, The *Normal EEG of the Waking Adult*, Electroencephalography 5$^{th}$ ed., Lippincott Williams & Wilkins.

[22]Nielsen, J. 1993, *Noncommand user interfaces*. *Commun. ACM, 36, 4 1993), 83-99.*

[23]Nijholt, A., Tan, D., Allison, B., Milan, J., and Graimann, B. 2008, *Braincomputer interfaces for HCI and game*', in CHI '08: CHI '08 extended abstracts on Human factors in computing systems, pp. 3925–3928, New York, NY, USA, (2008). ACM.

[24]Nijholt, A., Tan, D., Pfurtscheller, G., Brunner, C., Mill´an, J.R., Allison, B., Graimann, B., Popescu, F., Blankertz, B., and M¨uller, K.R. 2008, *Brain-Computer Interfacing for Intelligent Systems*, IEEE Intelligent Systems 2008, 23(3), 72–79.

[25]Neurosky 2012, *Neurosky,* viewed on 15 July 2012, <http://www.neurosky.com/>

[26]Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N. and Wolpaw, J. R. 2004, *BCI2000: A General-Purpose Brain-Computer Interface (BCI) System*, IEEE Transactions on Biomedical Engineering, vol. 51, no. 6, June 2004, viewed 24 March 2012, <http://www.bciresearch.org/paper.pdf>

[27]ThinkGeek 2012, *MindFlex Duel*, viewed on 5 April 2012, <http://www.thinkgeek.com/geek-kids/7-13-years/e813/>

[28]Thorisson, K. R., Koons, D. B. and Bolt, R. A. 1992, *Multi-modal natural dialogue*. In *Proceedings of the Proceedings of the SIGCHI conference on Human factors in computing systems (Monterey, California, United States, 1992).*

[29]Van Dam, A. 1997, *Post-WIMP User Interfaces, Communications of the ACM*, February 1997, Vol. 40, No. 2.

[30]Van Dam, A. 2000, Beyond WIMP. *Computer Graphics and Applications, IEEE, 20, 1 2000, 50-51.*

[31]Van Dam, A. 2001a, *User interfaces: disappearing, dissolving, and evolving*, *ACM, 44, 3 2001, 50-52.*

[32]Van Dam, *A. 2001b, Post-WIMP user interfaces: the human connection, Springer-Verlag London Ltd, Godalming.*

[33]Van de Laar, B., Bos, D. O., Reuderink, B., and Heylen, D. 2008, *Actual and Imagined Movement in BCI Gaming*, viewed 24 March 2012, <http://www.aisb.org.uk/convention/aisb09/Proceedings/AIGAMES/FILES/vandeLaarB.pdf>

[34]Van Erp, J., Lotte, F., Tangermannm, M. 2011, *Brain-Computer Interfaces for Non-Medical Applications: How to Move Forward*, IEEE 2011.

[35]Wang, Q., Sourina, O., Nguyen, M.K. 2010, *EEG-based " Serious" Games Design for Medical Applications*, 2010 International Conference on Cyberworlds.

[36]Wolpaw, J. R., Birbaumer, N., McFarland, Pfurtscheller, G. and Vaughan, T. M. 2006, *Brain-computer interfaces for communication and control*, viewed 24 March 2012, <http://www.ai.rug.nl/~lambert/projects/BCI/literature/serious/non-invasive/BCI-for-communication-and-control.pdf>

# Appendix A: Questionnaire

The following pages show the questionnaire as it was presented to participants. The NASA task load index component of the questionnaire was completed separately on a computer using an application that automatically calculates the total workload.

Questionnaire

1. Please select the age range you fall into:

   ☐ Under 21

   ☐ 21 to 30

   ☐ 31 to 40

   ☐ 41 to 50

   ☐ Over 50

2. Please select your gender:

   ☐ Male

   ☐ Female

3. Please select the category that best describes you:

   ☐ Student

   ☐ Academic Staff

   ☐ Other

4. If in the previous question you chose "Student", please state the course you are studying:

   ________________________________________

   ________________________________________

5. On average how many hours a week would you spend using a computer:

   ☐ Less than 10

   ☐ 11 to 20

   ☐ 21 to 30

   ☐ 31 to 40

   ☐ Over 41

6. On average how many hours a week would you spend playing games (inclusive of gaming on any consoles and the PC):

   ☐ Less than 10

   ☐ 11 to 20

   ☐ 21 to 30

   ☐ 31 to 40

   ☐ Over 41

7. Of the following game systems which have you used (select as many as is relevant):

   ☐ Xbox

   ☐ Playstation

   ☐ Wii

   ☐ Other (please specify): ________________________________

8. Of the following motion interaction technologies, which have you used (select as many as is relevant):

   ☐ Microsoft Kinect

   ☐ Playstation Move

   ☐ Wii Remote

   ☐ Other (please specify): ________________________________

9. Have you ever used any form of brain computer interface device (BCI)?

   ☐ Yes

   ☐ No

10. If you answered "Yes" to the previous question could you please specify any relevant details about what device you have used and in what context (eg, for other research testing? Was it game related?):

    ____________________________________________________________________________

    ____________________________________________________________________________

Please enter answers to the questions on this page in the boxes on the right.

11. On a scale of 1 to 10, how easy did you find it to use the input from the brain computer interface to complete the required tasks? (1 = very hard, 10 = very easy)

12. On a scale of 1 to 10, to what extent do you feel the extension of the brain computer interface improved the entertainment of gameplay? (1 = the BCI added nothing to the entertainment, 10 = the BCI added a lot to the entertainment)

13. On a scale of 1 to 10, to what extent do you feel the extension of the brain computer interface added more challenge to the tasks? (1 = no extra challenge, 10 = a lot of extra challenge)

14. On a scale of 1 to 10, how replayable do you feel games such as those that have just been played are? (1 = not replayable, 10 = would regularly replay)

15. On a scale of 1 to 10, how easy did you find it to change between calm and stressed states to complete gameplay tasks? (1 = very hard, 10 = very easy)

16. On a scale of 1 to 10, how easy was it to interact with the game using the physical movement provided by the Microsoft Kinect? (1 = very hard, 10 = very easy)

17. On a scale of 1 to 10, how easy was it to interact with the game using the speech input provided by the Microsoft Kinect? (1 = very hard, 10 = very easy)

18. On a scale of 1 to 10, how well do you feel the combination of the brain computer interface and the Microsoft Kinect provided a better gameplay experience? (1 = it made it a worse experience, 10 = it made it a much better experience)

19. On a scale of 1 to 10, how interested would you be in seeing brain computer interfaces used in games on the market? (1 = not interested, 10 = very interested)

20. On a scale of  to 10, how interested would you be in seeing games that use both the Microsoft Kinect and brain computer interfaces on the market? (1 = not interested, 10 = very interested)

21. Please write have any comments you would like to add relating to the scales in the previous questions here. (if you wish to clarify a reason for choosing a particular number)

22. How was your experience of using the brain computer interface?

23. How was your experience of using the Microsoft Kinect?

24. What are your opinions about combining the two input technologies?

Thank you for taking time to respond to this questionnaire.

## Appendix B: Information Supplied to Participants

Prior to being equipped with the Emotiv EPOC and the completion of the puzzles participants were provided with the information on the following pages. It provides details about the task the participants would be expected to complete and then allowed for any further questions to be asked prior to beginning each session.

**Dr Brett Wilkinson**
School of Computer Science, Engineering, and Mathematics
Faculty of Science and Engineering

Room 238, Information Science and Technology Building
Sturt Road, Bedford Park SA 5042

GPO Box 2100
Adelaide SA 5001
Tel: +61 8 8201 3662
Fax:+61 8 8201 2904
Email: brett.wilkinson@flinders.edu.au

Web www.flinders.edu.au/people/brett.wilkinson



# INFORMATION SHEET

**Title:** ‘Evaluating HCI using the Microsoft Kinect augmented with Non-invasive BCI’

**Investigators:**

Mr Peter Mitchell
School of Computer Science, Engineering and Mathematics
Flinders University
peter.mitchell@flinders.edu.au

Dr Brett Wilkinson
School of Computer Science, Engineering and Mathematics
Flinders University
Ph: 8201 3662
brett.wilkinson@flinders.edu.au

**Description of the study:**

This study is part of the project entitled ‘*Evaluating HCI using the Microsoft Kinect augmented with Non-invasive BCI*’. This project will investigate the usability and applicability of brain-computer interface devices to augment computer game interactions. This project is supported by Flinders University School of Computer Science, Engineering and Mathematics.

**Purpose of the study:**

This project aims to find out:

- If brain-computer interfaces are appropriate for game play;
- If a user is capable of maintaining appropriate relaxation levels to effectively engage with the game;
- If the use of natural user interfaces for game play impacts on the use of brain computer interfaces.

**What will I be asked to do?**

Participants will be asked to play a collection of short “mini” games that will explore the use of gestures and brain-computer interactions. These games will require controlled movements with the hands and select statements with voice to complete specific game

actions.  To augment this interaction the user will be asked to try and maintain a certain level of relaxation.  Depending on the game requirements this level of "calmness" will provide the player with more options within the game.
You will then be asked to complete a brief survey on your perceptions of the games and the control mechanisms used followed by an analysis of the workload required to use the interaction techniques.
The game session and follow up questionnaires are not expected to take more than one (1) hour.
This is voluntary.

**What benefit will I gain from being involved in this study?**
The sharing of your experiences will improve the planning and delivery of future programs. There will be no direct benefit to you as an individual for taking part in this evaluation.

**Will I be identifiable by being involved in this study?**
We do not need your name and you will be completely anonymous. All data collected for the project will be de-identified, and any comments you make will not be linked directly to you. However, we will be utilising the video stream captured by the Microsoft Kinect device to analyse movement and game specific interactions.  Please take the time to consider this issue and ask any questions about the use of the content.

**Are there any risks or discomforts if I am involved?**
The investigator anticipates few risks from your involvement in this study. We only seek to evaluate your experiences in using these different interaction techniques. If you have any concerns regarding anticipated or actual risks or discomforts, please raise them with the investigator.

**How do I agree to participate?**
You can agree to participate by return email.  Participation is voluntary. You may answer 'no comment' or choose not to answer any questions, and you are free to withdraw from the tests/questionnaire at any time without effect or consequences.

**Thank you for taking the time to read this information sheet and we hope that you will accept our invitation to be involved.**

*This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project No. 5633).  For more information regarding ethical approval of the project the Executive Officer of the Committee can be contacted by telephone on 8201 3116, by fax on 8201 2035 or by email human.researchethics@flinders.edu.au*

## Appendix C: Back Story

The puzzles in this project were not designed entirely randomly. The Tile Puzzle was the first puzzle that was designed. From this puzzle a back story was created to explain its purpose that could then be applied to the other puzzles to create a sense of immersion in the story. The idea for the first puzzle is that you are an individual who has been surveying a tomb located somewhere in the Middle East. The general theme of the game deliberately suggests that this may be Egypt. If you look at the world map seen in Figure 38 as it was shown earlier. There is a triangle shown at a point resting over Egypt and there are very finely written coordinates that do map to the real world location represented. It is unlikely that anyone would notice this particular level of detail certainly during testing as the backstory was less of a focus than the how to play aspect. It is a hidden feature that could be used in a future stage of the story.

Figure 38. World Map

Figure 39 then shows the back story information that is presented as an informational screen prior to the start of the first puzzle. It creates the entry point for beginning the first puzzle.

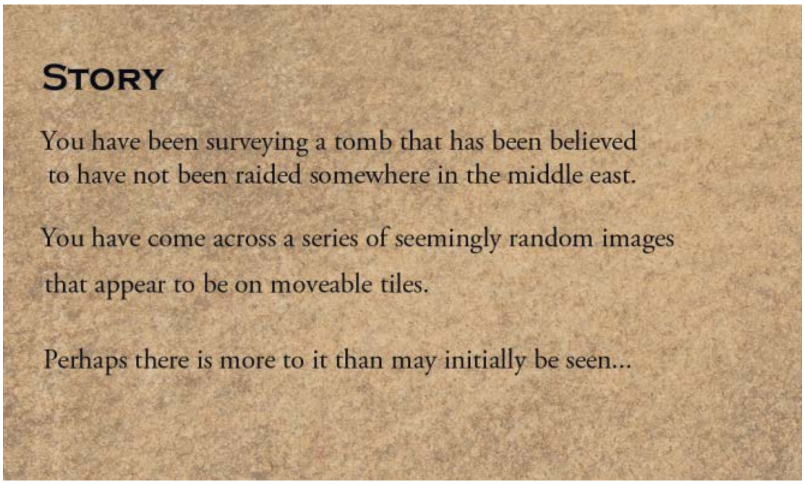



Figure 39. Tile Puzzle Back Story

Figure 40 shows the back story for the Street Puzzle as seen in game prior to the puzzle. The back story is based on the idea that nothing has been found at the site where you have been digging. It is determined that the tomb has recently been emptied. The shadowy figure seen running away you follow back to the market place. Then the puzzle for this level is to follow the thief back to their hideout without being detected so that you can discover where the treasure is being held.

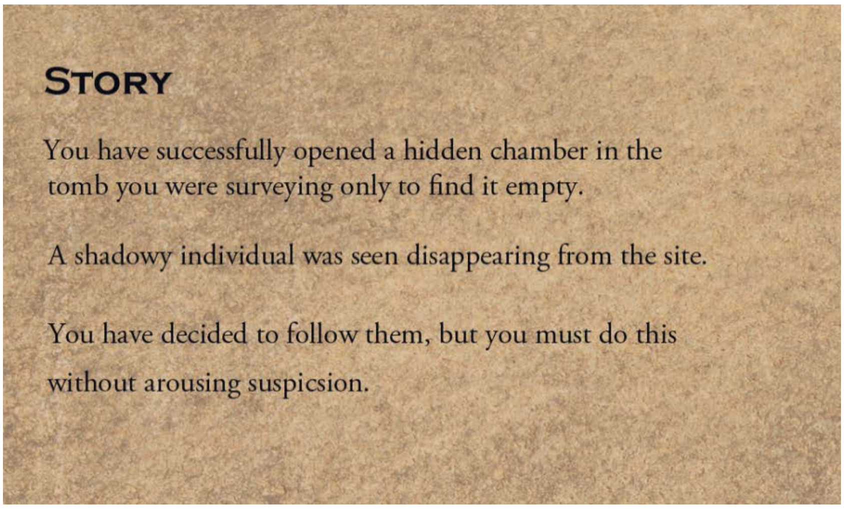



Figure 40. Street Puzzle Back Story

Figure 41shows the back story for the River Puzzle as seen prior to the start of the puzzle. In this final puzzle the story is based on having successfully followed the thief. You have confronted them and in a bid to escape they have thrown the treasure into the river. Then instead of chasing the thief you decide you must save as much of the treasure as possible. So the treasure must be grabbed as quickly as possible so that as much of it can be saved possible. The time limit used in the level represents the time it takes for the last of the pieces to be floating past you moving any possibility of further objects you want floating past.

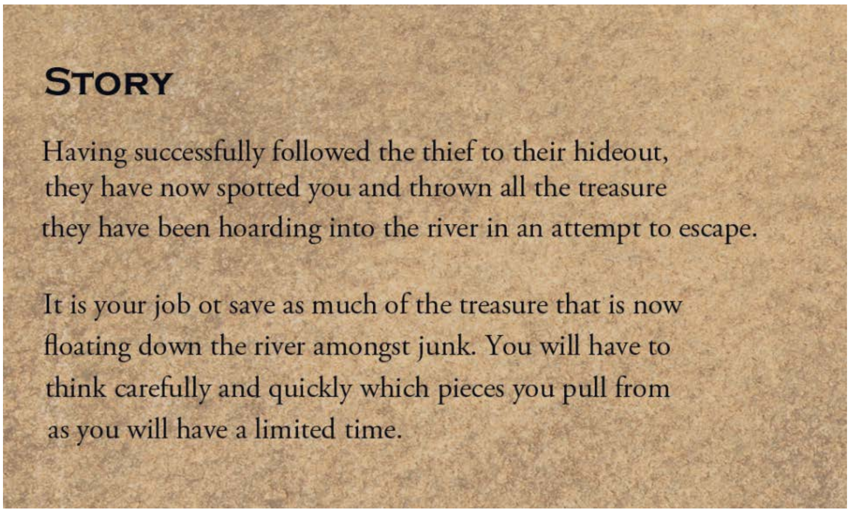



Figure 41. River Puzzle Back Story

This set of simple puzzles could be complemented with a continued story that could go in a number of ways. One of the most obvious would be to have the player try and track down the tomb thieves to confront them again. This is just one of many ways this game could be continued.