EEG Emotion Recognition in Videogame Play

José Rodríguez

Master of Science in Medialogy

Supervisors: Michael Boelstoft Holte, Aalborg University
Søren Fridmodt-Møller, Aalborg University
Lennart Nacke, University of Waterloo

Abstract

This study explores emotion recognition in videogames using electroencephalographic (EEG) data. Presently, emotion recognition using pattern recognition techniques has not yet been investigated in videogame play. This research is motivated by the possibility of retrieving insights into player experience from EEG signal during gameplay, which aims to contribute to Games User Research as an emerging discipline in the study of videogame design and their interaction with the players. In order to investigate emotion recognition several approaches for feature vector creation and classification algorithms were employed in order to assess which combination offered higher accuracy in classification. A maximum of 33.48% of classification accuracy was achieved by the Nearest Mean Classifier in the classification of four different emotions. Such low results suggests the collection and pre-processing of data from a dynamic activity, such as videogame play demands novel approaches for filtering the EEG, rejecting of artifacts and selection of the emotional model into which map the EEG brainwave oscillations.

Acknowledgments

Firstly, I would like to thank my supervisors Michael Boelstoft Holte and Søren Fridmodt-Møller for their suggestions, corrections and comments along the course of this research. Their guidance has helped me to adequately express the ideas I hoped to communicate with this thesis.

Secondly I would like to thank Lennart Nacke, my external supervisor at the University of Waterloo in Canada. It is because of him I could travel to Canada and deepen my knowledge in Games User Research by joining the HCI Games Research Group at the Games Institute, an environment where I could discuss my thesis's ideas constructively.

I thank my labmate Rina Wehbe for being there in the moments of crisis, for listening and discussing my research ideas, and guiding me through the complex and tangled field of psychophysiology and electroencephalography.

Last but not least important I would like to thank my family and girlfriend for their support during this long journey; I could not have done this without them.

Table of contents

Introduction

Videogames are considered one of the preferred sources of entertainment. According to the Entertainment Software Association (ESA) annual report (ESA, 2015), 42% of the population in the United States play games regularly, generating a total of 15.4 billion dollars in sales. Player's average age is 35 years old and the average time spent playing is 3 hours per week. The development of videogames at the present time requires high investment of time and capital from indie developers to large videogame industry companies. As a result, it has become necessary for video game producers to ensure a return on investment. This scenario is where Games User Research (GUR) comes into place.

GUR is an emerging field that studies videogame design and their interaction with players; it feeds off Human Computer Interaction (HCI), Game Development and Experimental Psychology. To study this interaction, it uses different qualitative and quantitative techniques to gather insights into player experience. These methods feed mainly from players' self-reported information, from observers' interpretations of user behaviour or statistical analysis of user performance to allow researchers evaluate player experience. In industry, GUR main goal is to provide funded information that would contribute to the design of better games.

The concern for game user researchers to obtain insights based on objective and unbiased data has drawn attention to novel techniques that are sustained on psychophysiological data (i.e. physiological signals resulting from psychological states). These have been gaining attention from researchers due to their capabilities for objectivity and reliability as a source of user cognitive and emotional processes. Heart rate, electro-dermal activity, electromyography and encephalography (EEG) are some of the biometrics studied for games user researchers. However, due to the innovative and early character of these approaches, these techniques are not used solely in videogame evaluation.

EEG is the recording of electrical activity that originates from neural synapsis in the brain. Recent studies have shown that the stimulation from playing videogames has an associated impact on EEG. He, Yuang, Yang, Sheikholeslami &He (2008) found correlations between the impact of long term playing and task difficulty on the theta brainwave band. Salminen & Ravaja (2008) studied the effect of experiencing violent events in videogames on the theta band and alpha band frontal asymmetry (i.e. difference of activity between brain hemispheres). The impact of different level designs on EEG was researched by Nacke, Stellmach & Lindley (2011), finding that theta band average power is higher on level designs where the player is more immerse. Most of these correlation studies share in common analyses of the preponderance of certain EEG brainwave bands while participants are playing the videogame, often trying to find correlations between bands spectral power and player experience insights obtained from self-reported data such as questionnaires, interviews or behaviour analysis.

Pattern recognition techniques have been employed in the classification of user emotional states using EEG oscillations. Lin, Wang, Jung, Wu, Jeng &Duann (2010) achieved 82.29% of accuracy in the classification of four emotions (joy, anger, sadness, pleasure) in EEG resulting from music excerpt stimulation. Petrantronakis & Hadjileontiadis (2010) successfully classified six emotions (happiness, surprise, anger, fear, disgust and sadness) with 85.17% of classification accuracy. Murugappan, Nagarajan & Yaacob (2011), using EEG oscillations resulting from videoclip viewing managed to classify five emotions (happy, surprise, far, disgust and neutral) with 83.04% of accuracy. Different approaches to apply pattern recognition techniques are taken from the current literature in order to discern which ones can result in achieving the higher classification accuracy. Filtering techniques, emotion elicitation techniques, classification algorithms and feature extraction approaches represent the main takeaways from these classification studies.

This exploratory study is motivated by the possibility of creating a system for GUR that could be employed in order to retrieve user emotional states during videogame play. A tool used by game user researches to collect objective insights based on EEG oscillations would have great potential for uncovering game design issues. The capabilities of EEG to reflect player cognitive and emotional processes added to the ability of pattern recognition techniques to classify emotional states underpins the pursuance this Masters Thesis project. The research question that leads this research is:

How can pattern recognition techniques be applied to EEG data in order to reveal the emotional states of players during videogame play?

The expected contributions for this exploratory study are to investigate classification accuracy of emotional states and to uncover possible limitations in the classification of emotional states during videogame play; a stimulus not yet explored using pattern recognition techniques. For this purpose, the approach taken for carrying out the classification will be based on the findings of the current literature in emotion recognition.

The thesis structure is as follows. This chapter is a brief introduction to the study related research, problem and motivation of the study. Chapter 2, Background, discusses the main areas of research that sustains the study´s focus and techniques applied in the experiment. Chapter 3, Methodology, explains the approach taken for emotion recognition and the details of the experiment for the collection of data. Chapter 4, Analysis and Results, dissects the pre-processing techniques used for preparing the data and presents the results of emotion classification employing the collected data. Chapter 5, Discussion, offers an interpretation of the results and exposes the main limitations of the study.

Background

User research, neuroscience, psychophysiology, computer science, user experience and videogame design reflect in one way or another the different facets of this multidisciplinary Masters Thesis research project. This chapter represents the theoretical grounding that leads this project´s focus, experimental methodology, and analysis of findings. An overview of Games User Research (GUR) and its methods for gathering insights on player experience provides the motivation and ultimate goals for pursuing this research. This study aims to contribute to the current body of literature in GUR and aid in designing better games by gaining a better understanding of player experience.

Brain signals stimulated by videogame play and captured by electroencephalography (EEG) technology represents the raw material for this thesis. Basic EEG concepts and latest discoveries in the field of EEG related to videogames will be presented for gaining a clear understanding before diving into the project's methodology, analysis and presentation of findings. As a novel approach to emotion recognition research, the utilisation of pattern recognition techniques on video game stimuli is suggested. Key elements for understanding emotion classification problems are presented; these will guide this tentative proof of concept for retrieving emotional states during gameplay.

2.1 Games User Research

GUR is a young field that originates from the need of professionals in the videogame industry and academics to understand how players experience videogames. It is tightly bonded to HCI and feeds from disciplines such as psychology, ergonomics and user-centred design to devise evaluation methods. It is important to understand that GUR did not originate for conducting Quality Assurance (QA) tasks like difficulty levelling or spotting software defects (i.e. bugs) but for gaining a holistic understanding of the player's behaviour, motivations and affection. GUR is differently approached by the academic community and professionals in the videogame industry due to a difference in ultimate goals, time-constraints and to lessen scientific validation standards. In academia, the focus of studying the player in videogames is to create knowledge through rigorous scientific method aiming to understand better different aspects of the players' interaction with videogames. In the videogame industry, GUR steers towards finding validation to design decisions. A game user researcher works to offer production teams insights gathered at various stages of the development cycle (Nacke, 2015) that could lead to improvement of the current videogame design. This is an iterative process (McAlister & White, 2015) that requires identification of issues and validation of actions taken to fix corrected flaws. The influx of professionals with academic background and the utilisation of evaluation methods borrowed from HCI in videogame companies have inspired GUR in industry to employ the scientific method and formulation of hypothesis for later validation (Nacke, 2015). The dialogue between GUR teams and productions teams (i.e. software developers, artists, writers) is key and often potentially frustrating due to the relatively recent integration of GUR along the development cycle for many videogame companies.

Bernhaupt (2015) devises a framework for the classification of evaluation methods used in GUR according to the different approaches of the procedures used for evaluation and gathering User Experience (UX) data. This classification distinguishes between Expert-oriented methods (1); based on design principles and heuristics evaluation, Automated methods (2); works with data mining techniques to infer player behaviour and User-oriented methods (3); feeds from data collected directly from the player. This subchapter aims to describe briefly the different techniques nested within the previously described classification.

Expert-oriented methods

Videogame expert reviews involve the evaluation of several aspects of the videogame design such as interface, narrative and game mechanics. The main goal is to minimize usability issues and to ensure the game's playability and enjoyment; factors that have an impact on player experience (Sweetser & Wyeth, 2005). It is important to notice the influence of techniques directed to the evaluation of production software (Nielsen, 1994), which in several cases have been adapted to suit videogame evaluation. For these evaluations, design principles are used to evaluate several aspects of the videogames (e.g., consistency of interface design); these methods are described as heuristics. Using heuristics evaluation provides a practical, effective and low-cost method for evaluation, which could be implemented early at the development cycle (Hochleitner, Hochleitner, Graf & Tscheligi, 2015); however the expert's review can be subject to personal bias (Nacke 2015). For further information on heuristic principles and frameworks for videogame evaluation I refer to the work of Hochleitner et al. (2015), where a detailed description of 49 heuristics is provided.

Automated Methods

The collection of big data stemming from the interaction between videogames and players represents raw material for Game Analytics. Server logs on micropayments, in-game player behaviour, software bugs, player's preference and rejections amongst other sources of data can offer insightful information for understanding better the players and for improving the game. The use of Game Analytics is becoming a common practice on videogame industry, stressing mobile and social online gaming (Drachen, 2015).

The collection of data logged on game servers regarding how players interact with every feature of the game is coined as behavioural telemetry data (Drachen, El-Nasr & Canossa, 2013). Vast databases populated by behavioural telemetry data can be registered in real time; capturing myriad actions such as the place and time of a player's death, or which weapons are less used on a first person shooter. The collected data per se, does not provide any insights. It is after applying data mining techniques (e.g. classification, clustering, regression, etc.) when the data is translated into interpretable data called Game Metrics (Drachen, 2015). Game Metrics can offer developers hints on design decision before and after the game is launched (Wallner & Kriglstein, 2015). Nevertheless, inferences based on Game Metrics on players' experiences are susceptible of misinterpretations, such as hypothetically missing a cofounding variable when defining trends. Furthermore, Game Metrics can be offered to players themselves to learn about their own performance (e.g. number of kills and shooting accuracy score). Presenting Game Metrics to developers in a meaningful way is key for making the most of the analysis of the data. Visualisation techniques (e.g. heat maps, charts, etc.) that allow for quick understanding, interpretation and exploration of the data is vital for taking well-founded design decision (Wallner & Kriglstein, 2013).

User-oriented methods

Research methods used for the collection of data directly from the players will be classified according to the scope of this thesis as Traditional methods (1); this involves mainly qualitative data collected from the interaction between the researcher and the player and Physiological methods (2); quantitative data collected from players' physiological responses (i.e. data collected from body functions). Brief descriptions from these methods will be provided during the following sections, however not the entirety of them will be covered.

Traditional methods

Most videogame evaluations in academia, and especially in the videogame industry, have been using the following traditional methods. They are considered traditional due to their wide acceptance and long standing in videogame evaluations.

Think-aloud protocol

This method introduced by Lewis & Mack (1982) consists of having the tester or player verbalise their thought process while they are performing a set of actions predefined by the researcher or just exploring the game. This technique is widely being used in user experience evaluations and helps to expose usability issues (e.g. a cumbersome navigation through the videogame menu interface).The think-aloud protocol is usually carried out in a lab-controlled environment in the presence of a researcher who takes notes and facilitates the evaluation exercise.

Interviews

Interviews are mainly used early at the beginning of the development cycle (Bernhaupt, 2015) for gathering insights from the target players. During the concept or prototyping stages of game development it is important to evaluate opinions and initial reactions from a game while substantial changes can still be made than in advanced stages in the developmental cycle (e.g. Alpha and Beta versions). Transcribing and coding interviews is a time consuming process, nevertheless it helps to identify common trends coming from target players.

Focus groups

A focus group takes advantage of the interaction within group members for developing insights through sharing opinions and experiences (Kitzinger, 1995). This qualitative method consists in having a group discussion with members that represent the target players and a facilitator that guides and prompts the discussion towards the previously defined topic of discussion. It is used early in game development during the concept and prototyping stages. This technique is based on the assumption that through interacting with the other members of the group, meaningful opinions can be developed solely from that social interaction (Preece, Sharp & Rogers, 2015). Focus groups is an inexpensive method that allows conclusions to be drawn relatively quickly.

Behavioural analysis

This technique consists on observing the player's interaction with the videogame (Nacke, 2015). The facilitator could guide the player's action by predefining several tasks to do in the game and observing the player performance, or let the player explore a game level or map freely. Behavioural analysis is a technique commonly employed during playtesting sessions to uncover usability issues; it can be combined with other methods such as the think-aloud protocol or questionnaires. The figure of the facilitator is essential for interpreting the facts that occur without personal bias, moreover, active questioning during the event of difficult interpretation (e.g. the player sighs after performing a task) could help unveil issues that otherwise would remain uncover.

Questionnaires

Using questionnaires allows collecting self-reported data systematically to quickly find trends over a sample (Preece et al., 2015). It could be employed during or after the course of gameplay. It is usually conceived as a separate piece of evaluation however it can also be embedded into the videogame; appearing after certain events for gathering fresh impressions over game features. Questionnaires can be used for evaluating myriad aspects of a game (e.g. narrative, aesthetics, functionality, etc.) or the game experience itself. There are standardised questionnaires specialised in measuring several aspects of how a player experience a videogame. For example, the GEQ questionnaire measures engagement in games (Brockmyer, Fox, Curtiss, McBroom, Burkhart & Pidruzny (2009); the ITQ questionnaire is designed to measure immersion (Witmer & Singer, 1998) or the PIFF questionnaire that evaluated presence, involvement and flow (Takatalo, Häkkinen, & Nyman, 2015).

Physiological methods

Evaluation methods that feed from body responses, such as body temperature, pupillary response, EEG and heart rate amongst others, to offer insights from player experience are known as physiological methods or biometrics. There is a growing interest on studying and decoding psychological states from physiological responses (psychophysiology), not only for the evaluation of videogames but for HCI regarding adaptive computer systems (Fairclough, 2009). There are several benefits that come from using biometrics, such as the instantaneous and unbiased character of the collected data (Nacke, 2013), however other challenges arise from interpreting and translating that information into usable metrics for evaluating player experience. Electro-dermal activity (EDA) and facial Electromyography (EMG) will be covered in this research. A deeper look into psychophysiology and EEG will be given in the following section.

Facial electromyography

This technique measures tension or electrical activity in a muscle of interest. Facial electromyography in GUR is mostly concerned with the corrugator supercilii and the zygomaticus major muscles as indicators of positive and negative emotions (Nacke, 2010; Hazlett, 2006). Tensed and relaxed states of these muscles will indicate when the player is smiling; zygomaticus major is tensed, or when the player is frowning; corrugator supercilii is tensioned. For facial EMG, electrodes are placed on the skin level and electrical activity, after signal processing is measured in microvolts (μV). For a deeper look at this procedure used for GUR I refer the works of Mandryk, Inkpen & Calvert (2006) and Ravaja, Saari, Salminen, Laarni, & Kallinen (2006).

Electro-dermal activity

Eccrine sweat glands are responsible for producing sweat in the human body. Emotional states can influence sweat production (Cacioppo, Tassinary & Bernston, 2007). EDA measures the changes of electrical conductivity on the skin capturing decreasing or increasing levels produced by a stimulus (i.e. a videogame). EDA is measured by two electrodes that can be placed at the fingers or palms. It is an inexpensive technique, non-intrusive and is relatively easy to interpret its results due to a high correlation between EDA levels and system arousal (i.e. being excited by a stimulus) (Lang, 1995). The studies of Mirza-Babaei, Long, Foley & McAllister (2011) and Drachen, Nacke, Yannakakis & Pedersen (2010) give a detailed explanation on how EDA is employed for videogame evaluation.

2.2 EEG and video games

Electroencephalography (EEG) is the record of brain oscillations provoked by populations of neurons electric fields and captured by a sensor placed on the scalp. The first human brain oscillations were recorded by the German psychiatrist Hans Berger in 1924 (Haas, 2003), however due to the novelty and scepticism of such discovery for showcasing insights of brain activity, it was not until 1934 when Adrian & Mathews (1934) corroborated his observations that EEG was considered as a valid tool for studying brain dynamics. The introduction of computers for data analysis introduced a big step in its development, allowing for the performing powerful computational techniques for EEG signal processing and analysis. With the release of the first *Neurosky* product in 2007 (“What is BCI,” 2015), EEG technology took a leap from the secluded medical research field to become a consumer product. Since then, EEG technology has become affordable in comparison to other technologies for brain imaging, such as fMRI, PET or CT, offering the public the possibility to explore brain dynamics for either research purposes or as a Brain Computer Interface (BCI).

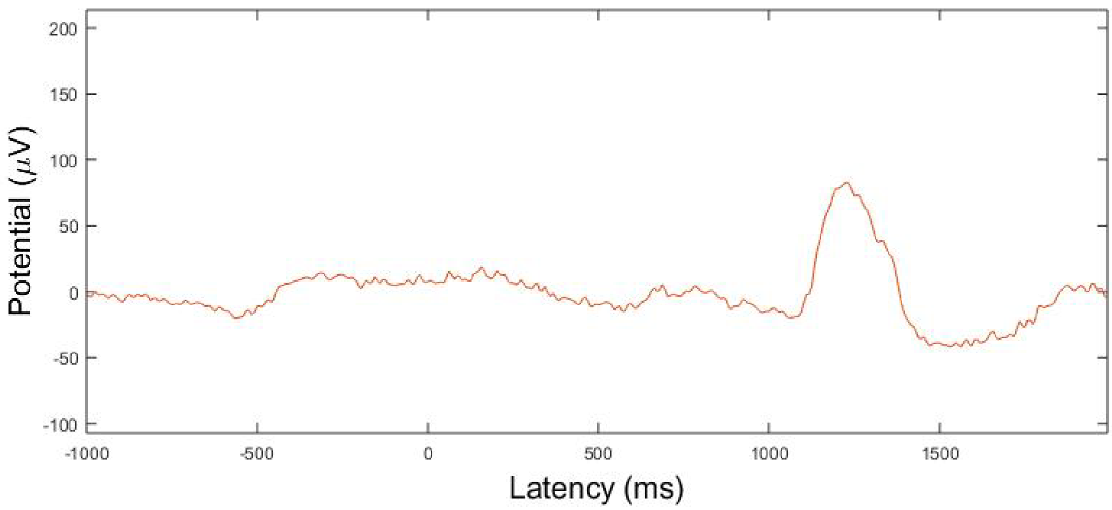


Figure 1: Raw EEG reading from a frontal channel

<u>EEG Basics</u>

Synapses between populations of neurons at the brain cortex produce electric fields that are captured by sensors or electrodes. What these electrodes read is the difference of voltage between the electrode of interest and the reference electrode. Basic plotting of raw EEG obtained by a pair or electrodes is characterized mainly by two variables; time and microvolts (μV) and brain oscillations featuring the change of voltage as a function of time(see figure 1).

Electrode placement

Multichannel EEG, recording with several electrodes attached to the scalp, affords a spatial dimension to EEG analysis that allows for the recognition of which brain structures could be involved during brain processes. It became clear to the scientific community that in order to compare and reproduce studies with precision and accuracy it was necessary to standardise electrode placement positions along the scalp. The 10-20 system proposed by Jasper in 1958 features the position of 21 electrodes (see figure 2) and is currently the most internationally accepted. Additionally, in order to obtain a higher spatial resolution, the original layout has been extended with new placements that would fill the intermediate spaces according to the 10-10 system (Chatrian, Lettich & Nelson, 1985) and the 10-5 system (Oostenveld & Praamstra, 2001) featuring a total of 128 brain electrode positions.

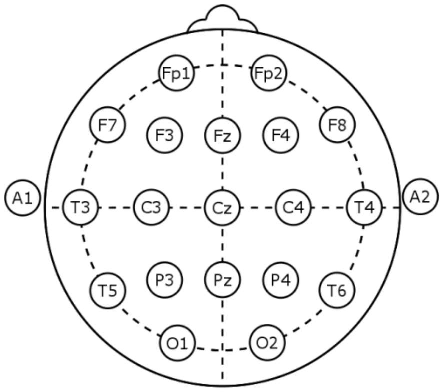


Figure 2: 10-20 system of electrode placement

Noise and artifacts

Before performing any EEG analysis it is necessary to filter the signal from different sources of contamination of non-cephalic signal. EEG is highly sensitive to external noise provoked by electrical equipment, poor electrode contact, power lines and poor grounding (Cacioppo et al., 2007). Notch filters at 50/60 Hz are commonly used to reduce non-biological noise.

Blinks, head movements and eye movement originate artifacts; these noises of biological nature can be problematic when recording EEG because it does not reflect brain neural activity and it can potentially contaminate signal from several electrodes. Artifact detection can be performed manually; nevertheless it requires possessing a trained eye in spotting contaminated EEG. Semi-automated methods such as Independent Component Analysis (ICA) applied on raw EEG readings have also been proved to remove successfully artifacts resulting from blinks, eye movement and line noise (Jung, Makeig, Humphries, Lee, Mckeown, Iraqui & Sejnowski, 2000). ICA separates the source of artifacts on multichannel

EEG by linear decomposition; once selected and remove ICA components that presumably produce artifacts (e.g. blinks), a corrected artifact free EEG is generated.

EEG analyses

Raw EEG data can be analysed with different techniques depending on the focus of the study. Different techniques can offer different insight into how the brain processes internal and external information. For instance some are more suitable for analysing very short time periods, for source localisation or for intracranial flow of information. The following describes the main techniques related to this study and literature.

Frequency analysis

Frequency analysis is grounded on the assumption that sinusoidal waveforms are compounded of mixtures of several waveforms with different frequencies and amplitudes (Cacioppo et al., 2007). Performing the Fast Fourier transform (FFT) on EEG signal, results in obtaining the contribution or several frequency ranges to the main oscillation. FFT is not the only algorithm for frequency analysis, nonetheless is the mostly used in the current literature. The frequency ranges or bands of interest resulted from applying the FFT are categorized as delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz) and gamma (30-45 Hz). Results from frequency analysis will typically show the power (μV^2) from each of these bands for a given EEG signal. The power is usually computed as the square of the amplitude.

Time-frequency analysis

Based on frequency analysis, Hemispheric Frontal Alpha Asymmetry (HFAA) compares the differences in power between opposing frontal lobes. HFAA is considered to play an important part on emotional processing as a moderator of emotion (Coan & Allen, 2004). Davidson's model establishes a relationship between approach-withdrawal emotion and its effect on the frontal lobe (Davidson, Ekman, Saron, Senulis, & Friesen, 1992).

In contrast to frequency analysis, time-frequency analysis adds a temporal dimension to the analysis of an EEG segment. Frequency changes can be visualised (see figure 3) and inspected for frequency shifts that can occur over the time a task takes. In this sense it is possible to observe how frequencies behave during the course of a cognitive process. According to Cacioppo et al. (2007), time-frequency changes can be computed by various methods such as Short Time Fourier transforms and Wavelet analysis. Pointed out by Cohen (2014), disadvantages that come with time-frequency analysis are the decrease of temporal resolution due to the time-frequency decomposition and the lack of literature on time-frequency studies related to cognitive processes in comparison to other analysis based studies such as Event Related Potential (ERP) or frequency analysis.

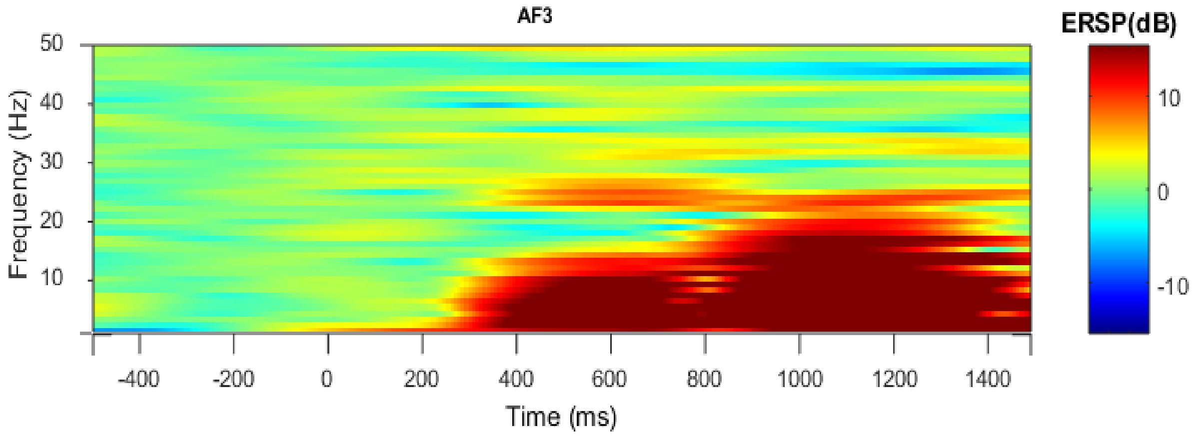


Figure 3: Time-frequency from electrode AF3

<u>Event related potentials</u>

The analysis of brain potentials time-locked to a stimulus is the basis of ERP. It focuses on observing waveforms components during exposure to a certain event that supposedly triggers cognitive processing. By observing these components during the stimulus exposure, covariations between them and event responses are targeted in order to identify how brain oscillations reflect cognitive processing. To obtain the ERP from an electrode position, the common procedure is to average all the samples to maximise the presence of brain signal related to the stimulus exposure and smooth the noise caused by background activity unrelated to event exposure (Cacioppo et al., 2007). As a result, a two dimensional plot voltage as a function of time (see figure 4) is presented and later analysed for the significant presence of a component such as the P300; a peak that appears at 300 milliseconds post visual stimulus.

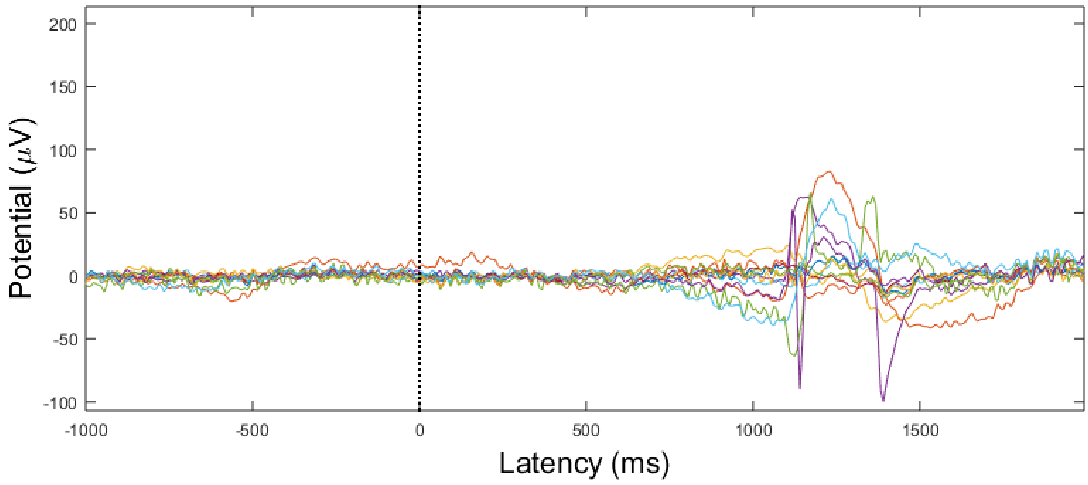


Figure 4: ERP from sample stimulus

EEG and other brain imaging techniques

EEG is not the only neuroimaging tool for studying the brain. Other technologies such as functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET), Computerised Tomography (CT) or Magnetoencephalography (MEG) can offer better results depending on what the subject of the study is. For instance, if the research focus entails the study of deep brain structures, functional localisation or the study of cognitive processes that unfold over long and uncertain periods of time (Cohen, 2014), EEG is not the most suitable technique. EEG has poor spatial resolution in comparison to the aforementioned techniques, which makes it less suitable for functional localisation. It can capture brain activity directly through voltage changes from clusters of neurons in the cortical areas; however it cannot be assessed from these measures exactly where the brain activity comes from due to Volume Conduction effects (Nunez, Srinivasan, Westdorp, Wijesinghe, Tucker, Silberstein & Cadush, 1997). The brain activity captured by electrodes is affected by myriad agents such as the subject's scalp, skull and brain folding. These singularities directly affect EEG readings, making it difficult to compute accurately the source of activity. On the bright side, EEG's outstanding temporal resolution has no rival when it comes to the study of neurocognitive processes. It can capture cognitive dynamics when they are happening with a resolution of a millisecond, and this makes it the most suitable technology to study emotional responses. Furthermore EEG headset devices have become exceptionally affordable in comparison with other technologies that are used mainly for medical diagnosis purposes, allowing researchers to operate low budget projects on EEG.

Psychophysiological EEG research on Videogames

The human nervous system is formed by the Central Nervous System (CNS) and the Peripheral Nervous System (PNS). The former's function is to integrate and coordinate the information that it receives from the whole body; it is the brain and the spinal cord. The latter's function is to connect the CNS to the rest of the body to receive information from around the body; it is divided into the Somatic Nervous System, which allows us to have control of our bodies; and the Autonomic Nervous System, which regulates body functions such as heart rate and pupil dilatation amongst others systems.

The field of psychophysiology studies the relationship between physiological phenomena provided by the CNS (i.e., neuron synapsis) and the PNS (e.g. heartrate, sweating, pupil dilatation, etc.), and psychological states that influence human behaviour and experience. It aims to understand how our bodies reflect psychological states in terms of physiological responses, for instance how a person's heartbeat increases in situations of danger or excitement, or how body temperature increases when somebody is blushing.

Linking psychological states to physiological responses is not a straight-forward process of Psychophysiological inference. As described by Cacioppo et al (2007) in *The handbook of Psychophysiology,* there are four possible types of relationship between psychological states and physiological responses.

- One-to-one: A psychological state is associated with only one physiological response.
- One-to-many: A psychological state is associated with a subset of physiological responses.

- Many-to-one: More than one psychological state is associated with a physiological response.
- Many-to-many: More than one psychological state is associated with a subset of physiological responses.

While the one-to-one-relationship would be the ideal for offering direct mapping between psychological states and physiological responses, it is more likely that the other types are the case, which increases the complexity of the psychophysiological inference. Physiological signals or responses offer objective and continuous data directly from the subject without being marked by the subject's interpretation, and without disturbing the gameplay (Fairclough, 2009). Nevertheless, as described by Nacke (2015), physiological responses tend to be volatile, variable and difficult to interpret without high level of experimental control.

Research on videogames with EEG

Several techniques are currently being used to study the players' experience through their physiological responses. These techniques (e.g., EDA, EMG, EEG, etc.) are not usually used alone but in combination with other psychophysiological or more traditional user research techniques, such as interviews, questionnaires or focus groups. For a brief description of physiological evaluation techniques read the previous section.

EEG psychophysiological research has traditionally been linked to medical or psychological purposes in order to offer insights into how the brain behaves and processes information under certain conditions. There is an extensive literature on correlations between frequency bands and psychological states. As an example, the delta band (1- 4 Hz) is found to be prominent during sleep (Cacioppo, 2007), theta activity (4-8 Hz) is associated with focused attention or drowsiness (Schacter, 1977) and alpha band is related to relaxation states (Cacioppo, 2007). These EEG correlates might help interpret the results from EEG analysis on videogames. They cannot however be considered as definitive references because they are highly dependent on the experimental conditions. Triangulating EEG results with other methods or design conditions may be the tenet for solid interpretation of results (Nacke, 2015).

Several studies have used videogames as stimuli for studying brain dynamics with EEG. The lens used by the academics in these studies set to consider the game as a source of stimuli for inducing psychological states that games could potentially cause. For instance, if the research focus revolves around studying how EEG potentials behave under situations of stress, a horror game with a high level of realism could induce stress and anxiety levels necessary for the collection of data

In this line of work, Schier (2000) studied drivers' attention by analysing power spectra from the alpha band (8 - 13 Hz) from two participants playing Need for Speed 3 (EA, 1998). This study was grounded on the findings of Schwartz, Salustri, Kaufman & Williamson (1989) that associated the suppression of alpha waves to the increase of attentional activity. Results from this small study with only two participants showed differences in alpha wave between driving and observation of driving tasks. He and his colleagues (2008) analysed spectral brain mapping of the theta band due to its association to mental workload (Gevins, Smith, McEvoy & Yu, 1997). For this purpose they compared spectral brainmaps between subjects playing strategy and competitive games that showed a substantial predominance of theta during the strategy game in comparison to the competitive one. Malik, Osman, Pauzi & Khairuddin (2012) investigated brain activation while playing videogames on large screens by means of spectral power, coherence and connectivity analysis. Experimental results showed more activation compared to baseline recording although proper experimental control was missing.

A different perspective is employed by academics in the GUR community, whose focus is on understanding how games may affect player's experiences by means of EEG. By gaining a better knowledge on the impact of certain game elements it would be possible to make well-funded design decision maximising player satisfaction. Nacke (2010) studied the impact of different interaction modes (Wiimote VS Controller) through the analysis of subjective ratings using the GEQ questionnaire (Brockmyer et al., 2009) and EEG spectral potentials from the alpha, beta, theta and delta bands. Results showed positive correlation between alpha power and negative rating in the GEQ, positive correlation between tension rating and alpha power on the controller and positive correlation between delta power and negative affect on the controller. The impact of level design (Nacke et al., 2011) was investigated by comparing brain potential to answers from the GEQ. The participants played three different levels conceived to make the player be in immersive, boredom and flow states. Correlation between immersion levels and theta band were founded. In-game learning was studied by Wehbe & Nacke (2013) by analysing the Mu rhythm (8-12 Hz) for its association with the mirror neuron system, and HFAA for its association for measuring arousal. Mu rhythm was found to be more active when the subjects were watching a person using the game than when they were playing for themselves, outlining the activity of the mirror neuron system for learning by modelling (Rizzolatti & Craighero, 2004). HFAA showed a higher level of arousal when playing the games in comparison to watching somebody play it.

In Salminen & Ravaja (2007), three types of event types were analysed by comparing brain oscillations through frequency analysis. By analysing oscillation changes from each of the events they found a statistical difference between each of the oscillations, confirming the impact of game events on brain dynamics with EEG. In Salminen & Ravaja (2008), they analysed violent events in a First Person Shooter (FPS) from EEG recordings using frequency analysis to find event related changes in brain potential; theta activation was found to be predominant during violent events.

Studies of videogames by means of EEG data have provided different frameworks and findings that have allowed correlating brain dynamics to different aspects of gaming. At present however, the sole use of this technique for assessing user experience of videogames is not considered sufficient for an adequate evaluation due to the difficulty of its interpretation and the lack of related literature specialised in videogame evaluation through EEG. Therefore it is necessary to use these techniques in conjunction with others in order to reveal a well-funded relationship between games aspects and psychophysiology (Nacke, 2015).

2.3 Pattern recognition techniques on EEG

Creating systems that would be able to discern between different emotional states from the user has been an important challenge in the development of emotional adaptive systems; potentially these could map the user emotions and deliver an adequate response to that form of HCI (Fairclough, 2009). The excellent temporal resolution from EEG, compared to other techniques to study brain dynamics, allows reading brain oscillation in real time from different locations on the scalp. Through experimental research on EEG, researchers (Davidson, 1992; Petrantonakis, & Hadjileontiadis, 2011; Park et al., 2011) have been able to associate these dynamics to planed events (i.e. presentation of stimuli) with the aims of finding a correlation between these and brainwave behaviour. In other words, it is certain that different stimuli can lead distinguishable brainwave oscillations (e.g. violent events can increase theta oscillations [Salminen&Ravaja, 2008]).

A pattern is a set of attributes that helps differentiating one object from another. In this research, the objects that are meant to be differentiated are emotions and the attributes that define these, are EEG potentials. Applying these techniques for emotion classification in EEG data represents an entangled problem due to the complexity of the brain dynamics.

Several attempts to classify EEG data into different emotional states have been made in the last few years of research on the field. In order to produce EEG recordings with a significant presence of emotional stimulation; typical stimuli modalities employed were music (Lin et al., 2010; Koelstra et al., 2011), pictures (Petrantronakis & Hadjileontiadis, 2010; Brown, Grundlehner & Penders, 2011) or videoclips (Murugappan et al., 2011; Wang, Nie & Lu, 2011). The innovative aspect of this study is the employment of videogame gameplay as stimuli for emotion elicitation and subsequent classification by applying pattern recognition techniques.

Retrieving emotional states from videogame gameplay could provide valuable insight into player experience to the gaming industry; as expressed by Järvinen (2008), the emotional response is a meaningful component of player experience. The possibility to map affective states into certain episodes of gameplay could help developers to understand how design decision can affect the player experience without interfering with natural gameplay due to EEG's capability to capture spontaneous and subconscious feedback of the user (Fairclough, 2009). Furthermore, apart from game evaluation purposes, other applications would involve a game outputting continuous adaptive response of the game to the players' affective state, changing in real time the game events to keep the player engaged, immersed, etc.

This subchapter on pattern recognition techniques applied to emotion recognition aims to devise the key elements, taking into account the need to solve pattern recognition classification problems and exemplify these with the latest studies in this area of specialization to ground this pattern recognition study.

Key elements in pattern recognition classification

This section elucidates the different questions that need to be tackled to successfully classify emotion from electroencephalographic data. These, will vary depending on the nature of the emotions meant to be classified and the stimuli chosen to elicit these emotions. Current literature reviews on this field by Konar, Halder & Chakraborty (2014) and Kim, Kim, Oh & Kim (2013) give a description of the different techniques used by researchers in the latest studies in the field of emotion recognition.

Emotional model

Different emotional models are considered when it comes to discerning between different emotions. It is not the purpose of this study to discuss what emotion is or which is the best procedure to categorise and measure them, however the main models will be exposed. Two main approaches to this problem are taken by current research in emotion recognition; envision emotions as discrete emotions or adopt a bi-dimensional model of emotion based on arousal and valence. Classification problems of emotions is mostly done under the model of discrete emotions proposed by Ekman (1987) who described basic emotions as happiness, surprise, anger, fear, disgust and sadness. In his latest work, he added several more complex emotions formed by the combination of the aforementioned ones (Ekman, 1999) .The bi-dimensional model of emotion was adapted from Russell & Mehrabian's (1977) work and classifies emotion according to different levels of arousal and valence (see figure 5). This categorisation is often made by using the Self-Assessment Maniking (SAM) questionnaire (Bradley & Lang, 1994). Valence defines an emotion as positive or negative (e.g., happiness is a positive emotion and sadness is a negative one) and arousal as the calmness or excitement that the emotions brings (e.g., surprise has a high arousal value).

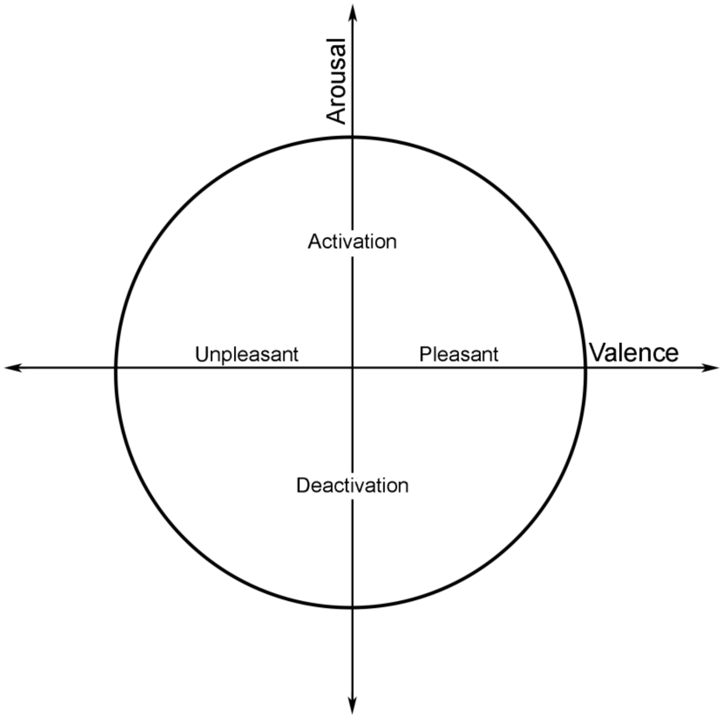


Figure 5: Bi-dimensional arousal-valence model of emotion

Emotion elicitation

In order two produce meaningful raw EEG data from cognitive and emotional processing, it is necessary to expose the subjects to stimuli that can produce relevant emotional signals to use it for tackling the classification problem. The sort of stimuli used is usually pictures, audio-visual clips, music or sounds and the duration of the elicitation is normally being constricted to short periods of time (e.g., 5, 10, 20 or 40 seconds). In order to label the EEG data collected, taking into account the subjectivity of experiencing emotions from same stimuli, self-reported data from the subjects is necessary and usually pursued by the use of questionnaires such as the SAM (Bradley & Lang, 1994) or custom questionnaires where the subject chooses from a short set of emotions (e.g. happiness, sadness or fear). Several studies (Brown et al., 2011; Petrantonakis & Hadjileontiadis, 2010; Konstantinidis, Frantzidis, Pappas & Bamidis, 2012) have used standardised databases of pictures such as the International Affective Picture System (IAPS) (Lang, Bradley & Cuthbert, 2008). The IAPS offers a large collection of pictures with validated scores on arousal and valence used for obtaining EEG data with specific arousal and valence values. During the experimental phases of the previous described studies, the subject remains passive while being exposed to the stimuli. In this research the subject was actively playing a videogame, increasing the complexity of classification due to noise from muscle contractions, head and eye movement.

Feature extraction and selection

Data collected during emotion elicitation needs to be processed for the pursuance of meaningful features than can help discriminate different emotions with better accuracy. Extracted EEG data parameters depend on the techniques used to pre-process (i.e. preparing the data by extracting artifacts or removing baseline EEG activity) and process the data before creating features vectors. A feature vector is a vector that carries the information that describes one of the emotional events measured during the emotion elicitation process. They can vary in length and belong to a certain class or category. For extracting feature vectors from raw data there are different approaches used such as working with frequency powers (Lin et al., 2010; Brown et al., 2011) from different brainwave bands (i.e. delta, beta, alpha,theta and gamma), amplitude of ERP (Konstantinidis et al., 2012), wavelet transform (Murugappan et al., 2010), time-frequency analysis (Hadjidimitriou & Hadjileontiadis, 2012). Additionally, how the brain isthought to process emotional information, can affect in what way these feature vectors are being selected. For instance, several studies (Petrantonakis & Hadjileontiadis, 2010; Hadjidimitriou & Hadjileontiadis, 2012; Lin et al., 2010) have underpinned their feature selection based on hemispheric asymmetry, an analysis approach that states that difference of activation between hemispheres can be used to measure emotions (Davidson et al., 2004), predominance of frontal and frontal central lobes to elicit emotions and the predominance of alpha and beta bands for emotion activation elicitation. Other analysis techniques can be based on connectivity and coherence between brain lobes during emotion processing or EEG mapping and space oriented analysis. In *EEG-based emotion recognition in music listening,* Lin et al., 2010 tested several configurations of feature vectors working with frequency powers at each of the electrodes for each of the brainwave bands. The researchers classified four different emotions: joy, anger, sadness and pleasure through 30 second music clips, obtaining best accuracy (82.29%) by feature vectors based on hemispheric asymmetry on frequency power. The authors exposed that the better result from the asymmetry approach could be related to the suppression of artifacts that contribute to both hemispheres and that they could be removed by differential power asymmetry approach.

Classification and Validation

Classification consists on identifying the belonging of an observation to a certain category from a set of categories (Van Der Heijden, Duin, Ridder & Tax, 2005). In order to solve a classification problem, it needs to be discerned which unlabelled EEG signals belong to which emotions. Therefore, it is necessary to train a classifier with feature vectors from videogame stimuli. A classifier holds the decision rule (i.e. classification algorithm) that maps an observation to a certain category. Classifiers are trained with datasets of observations from the stimuli meant to be classified. However, two different approaches can be taken; supervised and unsupervised learning. Supervised learning trains the classifiers with identified observations (i.e. categories are known), while on the other hand unsupervised learning or clustering (Van Der Heijden et al., 2005) assigns a category to the data based on similarity.

Studies on solving classification problems test several classifications algorithms to find which ones perform better with the inputted stimuli data. In offline emotion recognition classification (i.e. classification not in real time), accuracy is mostly considered as the best indicator performance. However in a different context such as online emotion classification (e.g. a system that aids learning according to your mood), where time constraints are vital, the processing time is considered an important performance metric.

In order to validate the effectiveness of a classifier and calculate performance metrics, validation methods are used (Van Der Heijden et al., 2005). Regular validation partitions the known data into two datasets: training dataset and testing dataset. The classifier is trained with the training dataset and performance is tested with the testing dataset where classes are being obviated. Incongruences in classification are verified by comparing classification results to the testing dataset (Van Der Heijden et al., 2005). In the case of scarcity of known data, cross-validation techniques can be used instead of partitioning known data into training and testing datasets which can be detrimental to the performance of the classifier. Essentially, cross-validation partitions the data into several subsets N, using one for testing the classifier and the rest for training N -1. Performance is calculated after averaging the results of training and validation with each of the N subsets.

Classification algorithms used in emotion recognition in related studies are related to the families of linear classifiers, support vector machines, quadratic classifiers, kernel estimation or neural networks, however other families of classification algorithms can be used for stimuli of a different nature (e.g. voice recognition, computer vision). The Linear Discriminant analysis (LDA), Quadratic Discrimant analysis (QDA), K-nearest neighbour (KNN) and Support Vector Machine (SVM) are the most used classifier algorithms, obtaining in most cases (Petrantonakis & Hadjileontiadis, 2010; Lin et al., 2010; Wang et al., 2011) the SVM, the highest accuracy score in emotion classification.

Methodology

Emotion recognition on EEG data is usually performed with brain oscillations from participants that have been exposed to music excerpts (Lin et al., 2010), pictures (Brown et al., 2011) or videoclips (Wang et al., 2011) while remaining still during the course of the exposure. Due to the inexistence of previous studies on emotion classification from videogame stimuli, the application of pattern recognition tools for maximising classification accuracy was investigated. This methodology describes the experiment from which the EEG and self-assessed emotion data was collected and the approach taken to apply pattern recognition tools to the collected data.

3.1 Experimental design

A within-subjects design was employed. Emotional self-assessed questionnaire responses and EEG potentials represented dependent variables. Emotional self-assessed questionnaire responses were obtained at different game events during the gameplay of a videogame. EEG oscillations were recorded during the entire gameplay process and EEG epochs (i.e. EEG excerpt) were extracted according to the analysed game events. Video footage from the participants was recorded in order to code facial expressions during gameplay for artifact detection (i.e. muscle movement that contaminates EEG) and videogame footage was saved during the entire gameplay for later selection and analysis of EEG epochs related to game events of interest.

Protocol

In laboratory conditions, the participants were exposed to the videogame stimuli in order to record brain oscillations from emotional responses evoked by game events during the gameplay. The duration of the experiment was approximately of 60 minutes; it is itemised as follows.

A. Introduction and questionnaire completing. Duration: 10 minutes.

The participant was introduced to the laboratory and received a detailed description of the full experimental procedure. After being informed they were asked to fill in a consent form and a demographic questionnaire regarding their age, gender and some questions regarding any previous experience with videogames and with the games used for the experiment.

B. Participant and headset setup. Duration: 15 minutes.

The participant was seated on a comfortable adjustable office chair positioned towards the PC. The EEG headset was placed making sure that the different electrode positions were placed according to the 10-20 system (Jasper,1958) and that the participant was comfortable wearing it. The optimal connectivity between the electrodes and the scalp was verified by using the *Emotiv Xavier TestBench* software before the EEG recording.

C. Pre-recording play time. Duration: 5 minutes.

The participant was introduced to the game and game controls and played to assure that any difficulty experienced by the participant was not due to a lack of knowledge of game dynamics and controls. If necessary some adjustment was made regarding the headset and participants' commodity.

D. Videogame test and emotional self-assessment questionnaire. Duration: 25 minutes.

Videogame test: EEG recordings and screen gameplay footage recordings started at this time by using *Emotiv Xavier TestBench* for the EEG data and the open source software *Open Broadcaster* for the gameplay footage. The participant played 15 minutes or until the completion of one level. After the time or level completion, EEG and gameplay footage recordings were stopped. Emotional self-assessment: Every two turns (i.e. player's turn and A.I. turn), the participant was prompted to fill in the *Affective questionnaire* (see appendix) for the following two turns.

E. Debriefing. Duration: 5 minutes.

The headset was removed and the participants were thanked for their participation.

Participants

Voluntary participants from both genders and over 18 years old were welcomed to take part in the experiment. To participate in the study, it was required to have computer game competence; have played videogames at least one time during the last month. Participants with psychiatric or neurologic conditions that could have an adverse impact on cognitive and emotional processes were excluded from the analysis. The participants were compensated with 10$ CAD for their participation. According to Kivikangas (2010), a good sample size for physiological research is 28 participants in order to obtain significant results transferable to the population. The completion of this experiment therefore targeted total of 28-30 participants.

Stimulus information

One PC game was be employed for the experiment; *Worms Revolution* (Team17, 2012). It is a tactical turn-based videogame, not too fast paced, features that allow a clear isolation of game events during the gameplay for the extraction of EEG epochs. The game is set on a 2.5D space where two teams of worms are equally distributed around the map geography. Both teams have an equal amount of worm fighters and weapons which should be used in the most efficient way to defeat the adversary. The diversity of movements, weapons, geography and selection of fighters during the player's turn, allows myriad tactical possibilities for each turn. Each turn is limited to a maximum of 30 seconds. Single player mode where the participant plays against the game A.I., was chosen for the test. The weapon selection was restricted to the bazooka and the map used remained the same for every participant to maximise consistency between participant's tests.

Measures

EEG was recorded from every participant during gameplay. Gameplay footage was recorded for epoch extraction in synchrony with the selected game events. The participant's face was recorded during the experiment to verify the presence of artifacts for the EEG analysis. The self-assessment emotional questionnaire was filled in every two turns for two turns during the gameplay.

EEG

EEG was obtained by the *Emotiv Epoc+* headset. The headset has 14 electrodes (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) distributed according to the 10-20 standard system of electrode placement. Data was transferred via Bluetooth to the PC and recorded by *Emotiv Xavier TestBench*, the software provided by *Emotiv*. Epoc+ records with 128 Hz of resolution a bandwidth of 0.2 – 45 Hz with notch filters at 50 Hz and 60 Hz to suppress noise from surrounding electrical devices.

Further signal processing was carried out using *EEGLAB* (Delorme & S Makeig , 2004), an open source *Matlab* (The MathWorks Inc., 2015) toolbox for processing electroencephalographic data. The EEG recording from each game was segmented in epochs according to the game events subjected to emotional self-assessment. After extraction, the epochs were visually inspected and cleaned for artifacts. A frequency analysis was performed by obtaining absolute power estimates (μV^2) using fast Fourier transform (FFT) for delta(1-4 Hz), theta(4-8 Hz), alpha(8-12 Hz), beta(10-30 Hz) and gamma(30-45 Hz).

EEG epochs were extracted, taking as a time reference the moment of impact from the player's and A.I.'s attacks. After EEG pre-processing (i.e. scoring markers, filtering and artifacting), each game event EEG epoch was analysed using *EEGLAB* on *Matlab*. For analysing EEG epochs, these ones are sectioned in three parts, using the moment of impact as reference (see figure 6). The first section duration is 1-second from before the moment of impact (1 second offset) and the following two sections are two 1-second sections after the moment of impact (2 seconds offset). For each section, frequency analysis will give mean power estimates (μV^2) for every band (delta, theta, alpha, beta and gamma).For a detailed description of the pre-processing steps, see the analysis section.

Figure 6: Explosion caused by the impact of a bazooka rocket

<u>Self-assessment emotional questionnaire</u>

An emotional self-assessment of game events offered a subjective evaluation of emotion in user experience; necessary for applying pattern recognition techniques for supervised learning classification. Based on Russell's Circumplex model (1980), the participant was prompted to take the SAM questionnaire (Bradley, Lang & Peter, 1994) to measure the arousal and valence dimensions of emotions. As a result, physiological data (EEG) was linked to self-reported emotional ratings. According to Russell (1980), emotions can be represented on a bi-dimensional space. The Valence dimension ranges from high positive to high negative and denotes intrinsic attractiveness or adversity. The Arousal dimension ranges from high intensity to low intensity and represents the physiological/psychological activation. Participants' self-assessments using the SAM questionnaire (Bradley, Lang & Peter, 1994) rated Valence and Arousal on a 9 – point scale.

Event types on gameplay footage

Gameplay footage was employed to locate excerpts of interest in the EEG data that were subjected to analysis. Furthermore, the researcher coded for event types by observing the gameplay footage. A total of 4 different types of events were categorised from these observations:

- The player attacks the adversary.
- The player fails to attack the adversary.
- The adversary attacks the player.
- The adversary fails to attack the player.

Event types do not represent the targeted data for emotion recognition, nevertheless pattern recognition was utilised to explore the capabilities of event types to generate EEG patterns for classification.

3.2 Applying pattern recognition to classify emotions

As emotion recognition based on EEG from videogame stimuli has not been yet investigated, it is proposed to explore videogame offline emotion recognition by taking into account different approaches for dataset creation, feature vector creation and classification processes along the pattern recognition techniques workflow. Following a factorial experimental design, different parameter combinations were explored in order to find the ones that exhibit the best accuracy in the classification of the proposed emotions.

Class creation

Results from the self-assessment emotional questionnaire categorise each of the resulting EEG oscillations extracted from the gameplay. Four different classes were devised according to the participant's ratings on the valence and arousal 9 – point scales. On a 9-point scale, values of 4 or less were considered low and values of 6 or more were considered as high. Middle values such as 5 resulted in discarding that EEG resulting oscillation from further analysis. Additionally, a new set of classes was defined according the event types.

- HAHV; for high arousal and high valence.
- HALV; for high arousal and low valence.
- LAHV; for low arousal and high valence.
- LALV; for low arousal and low valence.

Feature vector creation

To create the feature vectors to build the dataset for the training of algorithms, frequency domain techniques were employed. Two different approaches for feature vector creation were used:

-**Fast Fourier transform** (FFT): Applying the FFT, absolute spectral powers from each of the frequency bands were extracted for every electrode at each of the 3- second epochs. Local baseline (1 second offset) was extracted from the average of the remaining 2 seconds (2-second onset) forming a feature vectors space of 70(14 electrodes X 5 bands).

-**Asymmetry index(** (ASi). By applying the FFT to obtain spectral power for each of the bands at each of the electrodes and subtracting the right side hemisphere electrodes' spectral power to the left hemisphere electrodes' spectral power; forming a feature vector space of 35([7 right hemisphere electrodes X 5 bands]-[7 left hemisphere electrodes X 5 bands]) .

Classification and cross-validation

In order to discern emotions from future uncategorised EEG data, a learning algorithm or classifier is trained according to the features vectors belonging to the classes on a dataset. The following parametric and non-parametric classifiers were tested for emotion classification. Parametric classifiers are trained using the mean and variance of the analysed data; the Nearest Mean Classifier (NMC), Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) were employed. Non-parametric classifiers are not based on the estimation of variance and mean for classification, radial and polynomial Support Vector Machine (SVM) classifiers and K-Nearest Neighbour (KNN) classifier were considered for classification.

Once a classifier is trained, a mapping or model is generated. This model's purpose is to discern between different classes (i.e. emotions or event types) when it is fed by unlabelled data of the same nature it was trained with. The performance for each combination of class creation, feature vector creation and classifier was tested using 10-fold cross-validation for validation of results.

Analysis and Results

This chapter presents the results of the experiment´s later analysis using pattern recognition techniques. A detailed description of the pre-processing steps for preparing EEG data is offered, stressing the importance of artifacting (i.e. cleaning data) for obtaining high quality EEG data. Furthermore, participants' descriptive demographics and classification results are provided. The code used for pre-processing EEG data and running classifications can be found in the appendix chapter.

4.1 Summary of demographics

The following is a summary of the participants' responses to the demographics questionnaire filled up before the experiment was conducted. A total of 33 participants were tested, however data from 3 of the participants were discarded due to technical failures during the course of the experiment. From the resulting 30 participants, 8 were female and 22 were male. The main age group was from 20 to 24 years old (40%), followed by 18 to 19 (34%), 30 to 34 years old (13%), 25 to 29 (10%) and 35 to 40 (3%). The average number of hours spent playing videogames daily was 1.9 hours (SD=1.02). The preferred platform for playing videogames was the computer (83.3%), followed by *Nintendo Wii* (26,6%) and *Play Station 3*(20%). The most played genres amongst the participants were strategy (66,6 %), first person shooters (60%) and action games (56,6%). Regarding their experience at playing the videogame used for the experiment, 53 % of the participants did not have any experience playing any game from the *Worms* saga (Team17, 2012).

4.2 EEG pre-processing

The readings provided by the EEG headset, *Emotiv Epoc,* needed of further pre-processing in order to be prepared for analysing the results of the experiment. In this section it will be described chronologically the steps taken for this purpose.

A. Scoring markers

The continuous EEG recording for each of the participants needed to be split in order to extract the epochs of interest for this study. Secondary markers on the EEG recording were introduced manually by the researcher at the beginning of each of the player's turn during the course of the experiment. These secondary markers were used as a reference to calculate the primary markers. By revisiting the gameplay footage, primary markers were defined by calculating the time difference between the secondary marker and the moment of impact of the bazooka.

B. Filtering the data

After importing the EEG recording file to *Matlab's* (The MathWorks Inc., 2000) toolbox, *EEGLAB* (Delorme & S Makeig, 2004). Further filtering of the signal was necessary in order to eliminate undesired frequencies. A notch filter at 50/60 Hz was applied to remove line noise (i.e. external noise provoked by electrical equipment). Additionally a bandpass filter was applied to isolate frequencies from 1 Hz to 45 Hz.

C. Epoching the data

Using the primary markers scored during the first step of pre-processing, epochs were extracted from the continuous EEG data. Each of the markers represents the reference point, from which 1 second offset and 2-3 second onset were extracted as an epoch of 3 seconds of duration (see figure 7). The duration of each of the participant's gameplays varied according to their performance, hence the number of epochs extracted from each of the participants varied in consequence. An average of 12.2 (SD= 4.76) epochs were extracted from each participant.

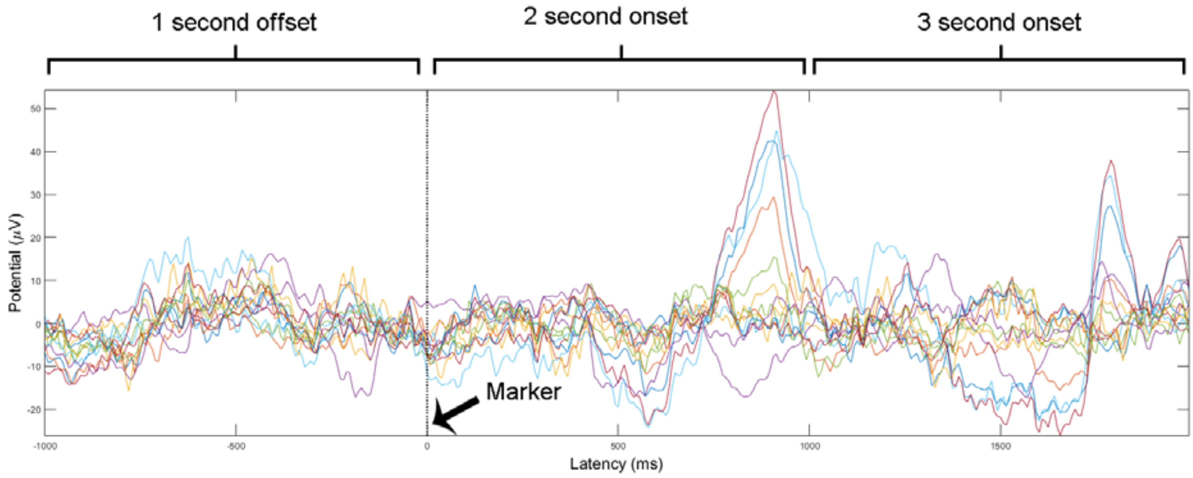


Figure 7: Example of epoch extraction

D. Cleaning the data

EEG data is highly susceptible to contamination from signal that does not originate from brain processes. Poor electrode connection, blinks, muscle movement or sweat can generate artifacts on the EEG readings that contaminate brainwave oscillations.

Each of the extracted epochs was first visually inspected to find paroxysmal noise (i.e. noise produced from sudden movement of headset). Epochs containing paroxysmal noise were rejected from further analysis. See figure 8.

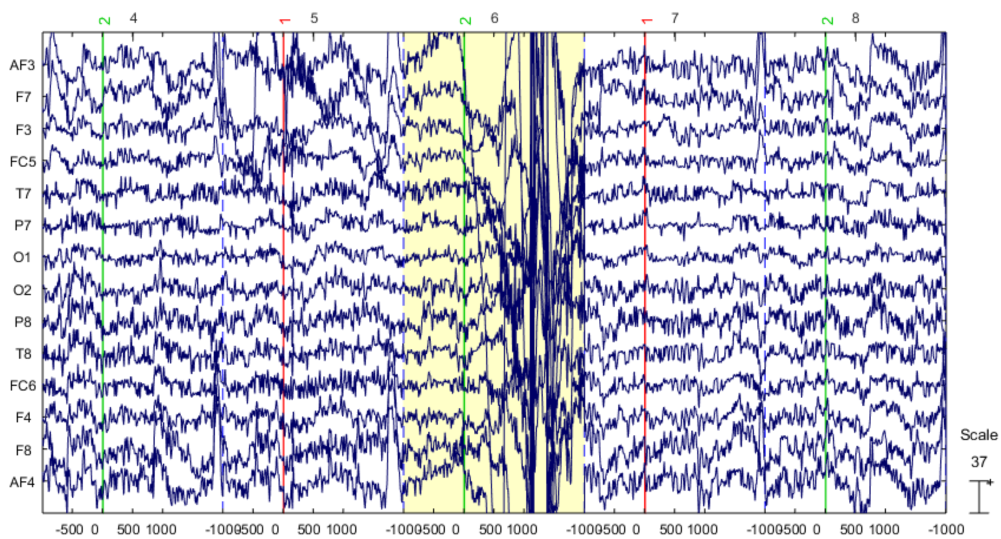


Figure 8: Massive paroxysmal noise in central epoch

Independent Component Analysis (ICA) decomposition is a validated technique widely used for eliminating artifacts (Makeig, Bell, Jung & Sejnowski, 1996). By applying ICA decomposition on multichannel EEG, it is possible to separate undesirable artifactual sources from EEG data. ICA decomposes EEG readings into different components and projects them on a scalp map, allowing for inspecting them and rejecting the ones that belong to artifactual sources. For instance in the following figures (see figure 9 and 10) it can be seen different components coming from ICA decomposition. Figure 9 represents a clear blink; as activity is focused only on the frontal electrodes (FC3 and FC4) and lack of activity in the rest of the electrodes. Figure 10 shows lateral eye movement as only frontal electrodes are mainly affected and the activity is polarised to one of sides. Once the undesirable components are removed, the total EEG readings are recomputed without the artifact sources.

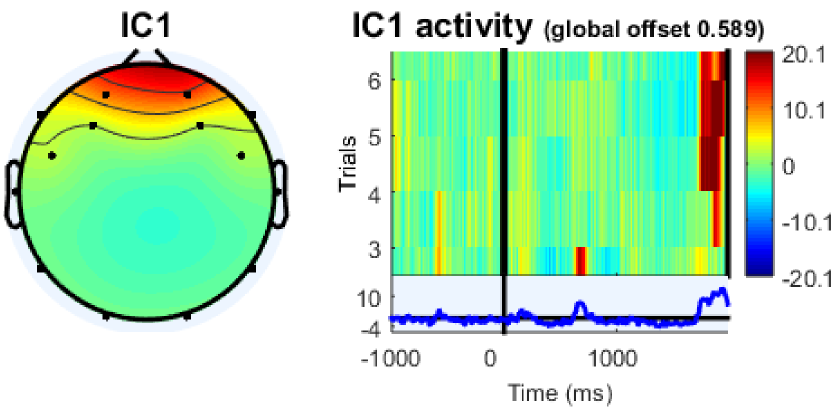


Figure 9: IC1 component showing artifact originated from blinking

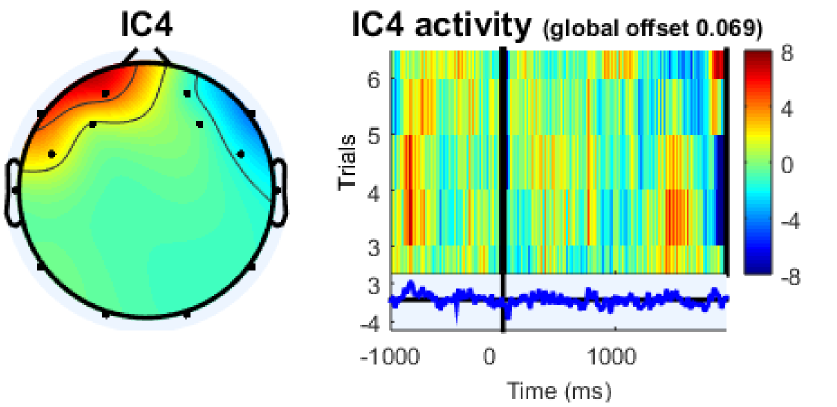


Figure 10: IC4 component showing artifact originated from lateral eye movement

E. Extracting spectral powers

After cleaning the EEG recordings signal, absolute power spectral values were extracted at each of the electrodes for each of the bands (delta, theta, alpha, beta and gamma) for each of the seconds of the extracted epochs. In order to obtain a resulting brain oscillation that reflected the exposure to the game event happened during the epoch, the absolute spectral values from the second pre-stimulus(1-second offset) were extracted from the average of the following post-stimulus seconds(2-seconds onset). The results from this calculation were used to build feature vectors that would form the datasets used for carrying out emotion recognition. Each of these epochs converted to a feature vector represents an instance in the dataset.

4.3 Classification results

Following the previous described process, five datasets were built in order to conduct further classification. The building of a dataset corresponds to the nature of the feature vector defined and the classes assigned to these feature vectors, which are described in the methodology section of this project. Two of the datasets (*ASi emotion* and *FFT emotion*) were built to verify if emotion recognition could be performed on the collected EEG data from videogame stimulation. Additionally, two more datasets (*ASi Event type* and *FFT Event type*) were created in order to verify whether using the same techniques the classification of event type would return high accuracy.

Every dataset had a total of four classes although the number of instances corresponding to each class was not equal across every class; hence the datasets were not balanced. Training a classifier with an imbalanced dataset results in achieving higher accuracy in classifying certain classes in detriment of the other classes with less appearance in the dataset. A fifth dataset, *Balanced ASi emotion*, was devised in order to envision to what degree classifiers would benefit from a balanced dataset. Nevertheless, as a consequence of balancing a dataset, the number of instances decreases as every class has to match the number of instances of the class with the least number of instances. The following table shows a relation between the datasets and classification accuracies achieved with parametric and non-parametric classifiers. Classifiers accuracies are averages of 10 runs of the classifier results using a 10-fold cross-validation technique. Calculations were performed with *PRtools* toolbox (Duin et al., 2004) on *Matlab*.

Table 1: Dataset description and results from emotion classification

Dataset description			Classification accuracy (%)					
			Parametric			Non-parametric		
Dataset	**No. & type of features**	**No. of instances**	**NMC**	**LDC**	**QDC**	**KNN**	**SVM-P**	**SVM-R**
ASi emotion	35 - ASi	216	**33.48%**	30.09%	29.62%	24.53%	29.01%	26.23%
FFT emotion	70 - FFT	216	**32.09%**	28.7%	19.59%	25.61%	30.09%	27.46%
ASi event type	35 - ASi	275	36.24%	36.96%	24.60%	**41.45%**	30.90%	36.96%
FFT event type	70 - FFT	275	38.80%	36.00%	14.35%	**39.53%**	33.33%	34.42%
Balanced ASi emotion	35 - ASi	168	**32.73%**	26.38%	29.56%	28.76%	25.79%	27.38%

Discussion and Limitations

The prediction of emotions or event types for each of the classifications performed on each of the datasets is significantly lower than what was expected based on related literature. In classifying emotions, the highest accuracy was achieved by the NMC classifier with a 33.48% using a feature vector based on hemisphere asymmetry. Regarding the classification of event types, the highest score was achieved by the KNN classifier with 41.45% employing hemisphere asymmetry feature vectors. Overall, the results for both classifications of emotions and event types are significantly low. If random chance in classifying one out of four classes is 25%, the mappings created by the classifiers are solely adding a maximum of 15% to event type classification and 8.48% to emotion classification. Therefore in any case these results could endorse the procedures described by this project's methodology to perform emotion classification.

Prediction accuracies from the emotion and event type datasets cannot be fully comparable. Firstly, both datasets have a different number of instances, which influences accuracy as the classifier benefits from learning the *nature* of a class. Secondly the oscillations from where absolute powers for each band were computed do not correspond between the datasets as different inclusion criteria was used for emotion categorisation and event type categorisation. It would therefore be risky to confirm that EEG oscillations with the applied methods represent the perceived phenomena (event type) better than subjective experience (emotion). As expected, balancing the *ASi emotion* dataset resulted in a proportional improvement of each of the classifiers' accuracy.

A comparison between feature vector creation in the emotion datasets and the event type datasets shows that *ASi* feature vectors perform better than the FFT ones. However, the differences in accuracies are significantly low as in any case is more than a 2%. Therefore, it cannot be verified, as Lin et al. (2010) and Petrantronakis et al. (2010) proposed, that feature vectors based on hemisphere asymmetry return higher accuracy than FFT feature vectors in emotion recognition. Classification accuracies across classifiers on a same dataset do not differ significantly. Surprisingly, in emotion recognition, the parametric classifier NMC returns the highest accuracy (33.48%), contradicting the findings of Petrantronakis et al.(2010), Lin et al. (2010), Frantzidis et al. (2010) and Chanel et al. (2009) where the non-parametric SVM classifiers returned the highest accuracy.

More than finding significant results, this exploratory study has evidenced the limitations that emerge when trying to perform emotion recognition in videogames. Judging from the accuracy results, the techniques and the findings obtained by the related literature, it is likely that the collection of data procedures for this study might have been the cause of such low prediction accuracies. In comparison to other studies in emotion recognition (Petrantronakis et al., 2010; Lin et al., 2010), the stimulus and collection of data is significantly different. A participant that is exposed to a piece of music or the view of a picture or video clip is passive to the stimulus as the observation of it is what is expected from the participant. When playing a videogame, the participant is actively engaged with the stimulus, which can be challenging for the researcher to have control over. For instance, some participants expressed themselves with facial expression, head movements, talking or gestures while playing the videogame even though they were advised to keep these to a minimum. The manner players naturally engage with videogames makes it difficult to obtain clean data without artifacts. Constraining players' interactions during videogame play to minimize artifacts on EEG data (e.g. using a neck brace to minimize movement) is also detrimental to the player experience and therefore nullifies this thesis's raison d'être. As an alternative, ICA decomposition was performed to eliminate the presence of artifacts, however this might have not been enough to obtain clean EEG data.

In comparison to related studies, the number of instances for each dataset, the balance of classes and the quality of the EEG headset (i.e. not a high quality medical EEG device) represents major drawback in this study. For emotion classification, 216 instances was the number of usable instances after cleaning the data. In contrast, Frantzdidis et al. (2010)

collected 4480 instances, Lin et al. (2010) 416 instances , Petrantronakis et al. (2010) 960 and Koelstra et al. (2011) 1280 instances for carrying out emotion classification. In the literature, the collection of data was carefully devised to produce highly balanced datasets. When exposing participants to pictures, music or video clips, to some degree it can be controlled which ones could elicit the desired emotions. Validated resources such as the IAPS (Lang et al., 2008) and IADS (Bradley & Lang, 1999) offer validated pictures or music that are mapped to a certain emotion. A videogame, unless designed for an experiment, cannot be fully controlled according to the researchers' purposes. The results of the player's interaction with the myriad game variables results in obtaining an unbalance number of instances per emotion/class.

Furthermore, it is possible that the emotion self-assessment procedure conducted during the experiment might have been too disrupting for the player's gameplay as they were prompted to stop playing and fill up a questionnaire every two turns. Embedding the emotion self-assessment within the game design would have minimized impact on the gameplay but increased the size of the study significantly as the stimulus design would have consumed high time resources. The mapping of emotions used, corresponded to the circumplex model of affect (Russell, 1980). Even though this validated model is used in related literature, additional models such as Ekman's (1999) or Plutchik's (1980) could be considered for future work.

Even though the results of this exploratory study have not been successful in classifying emotions, the research of videogames as new stimuli for emotion recognition identifies several challenges to be overcome in future studies. More specifically, devising new filtering techniques to minimize the presence of artifacts, designing non-disruptive emotion self-assessment collection methods and studying pattern recognition techniques represent the goals for achieving successful emotion recognition, which would contribute to generate valuable insight of player's experience.

Appendix

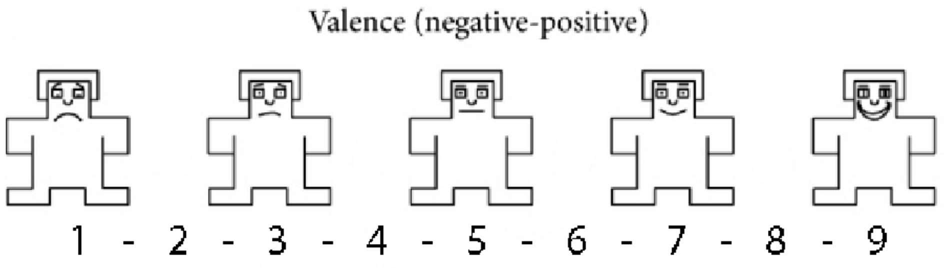
Questionnaires

Affective questionnaire

Please rate the last game event in terms of Valence and arousal.
What is valence and arousal?

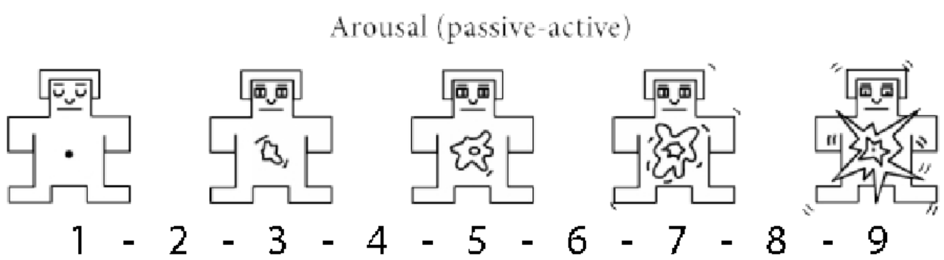
Valence refers to how positive or negative an event is.
Arousal reflects whether an event is exciting/agitating or calming/soothing.

*Required



*

1 2 3 4 5 6 7 8 9



*

1 2 3 4 5 6 7 8 9

Demographic questionnaire

*Required

Researcher Only- Participant ID *

Researcher Only- Participant ID

Choose one of the following answers *

- Male
- Female

Select your age group *

- 18 - 19
- 20 - 24
- 25 - 29
- 30 - 34
- 35 - 40
- 41 - 45
- 46 +
- Other:

What is your handedness? *

Choose one of the following answers

- Right handed
- Left handed
- ambidextrous

How long have you been playing videogames (years)? *

Choose one of the following answers

- 0
- 0 - 5
- 5 - 10
- 10 - 15
- 15 - 20
- 20 - 25
- 26+

How many hours of video games do you play per day? *

Choose one of the following answers

- 0
- 1
- 2
- 3
- 4
- 5
- 6+

What game consoles you own or play on a regular basis? *
Check any that apply

- Computer
- Play Station 1 (PS1)
- PS2
- PS3
- PS4
- PSP
- PS Vita
- Xbox
- Xbox 360
- Xbox one
- Nintendo Game Cube
- Nintento Wii
- Nintendo WiiU
- Ds
- Dsi
- DsXL
- 3DS
- Gameboy
- Gameboy color
- Other:

What genres of games do you play? *
Check any that apply

- Action
- Strategy
- First Person shooter
- Third person shooter
- Adventure
- Fighting
- MOBA
- MMO
- MMORPG

- [ ] Sports
- [ ] Party
- [ ] Puzzle
- [ ] Music/Dance
- [ ] Simulations
- [ ] Platform
- [ ] Racing
- [ ] Arcade style
- [ ] Mobile games
- [ ] Other:

Have you played any games from the Worms saga?

Choose one of the following answers

- ○ Yes
- ○ No

If yes, please rate your expertise from 1 to 5

1 being the lowest degree and 5 being the highest

1	2	3	4	5
○	○	○	○	○

Submit

Never submit passwords through Google Forms.

100%: You made it.

Code

EEGLAB pre-processing

```
%This piece of code does preprocessing, but not artifacting which will be
done semi-manually. If meant to be used with other files, change file names
in code.

%opening testbench file and retrieving data channels
[ALLEEG EEG CURRENTSET ALLCOM] = eeglab;
EEG = pop_biosig('C:\Users\Admin\Dropbox\Thesis\Data\testbenchfiles\ID26-
02.edf', 'channels',[3:16] );
[ALLEEG EEG CURRENTSET] = pop_newset(ALLEEG, EEG, 0,'gui','off');
EEG = eeg_checkset( EEG );
%assign channel locations
EEG = pop_editset(EEG, 'chanlocs',
'C:\\Users\\Admin\\Dropbox\\Thesis\\Implementation\\Raw eeg\\emotiv.ced');
[ALLEEG EEG] = eeg_store(ALLEEG, EEG, CURRENTSET);
EEG = eeg_checkset( EEG );
%importing events
EEG = pop_importevent( EEG,
'event','C:\\Users\\Admin\\Dropbox\\Thesis\\Data\\markers\\markersID26-
2.txt','fields',{'latency' 'type'},'timeunit',1);
[ALLEEG EEG] = eeg_store(ALLEEG, EEG, CURRENTSET);
%Notch filter 50 -60 hz
EEG = pop_eegfiltnew(EEG, 50, 60, 212, 1, [], 1);
[ALLEEG EEG CURRENTSET] = pop_newset(ALLEEG, EEG, 1,'gui','off');
%Bandpass filter 1hz
EEG = pop_eegfiltnew(EEG, [], 1, 424, true, [], 1);
[ALLEEG EEG CURRENTSET] = pop_newset(ALLEEG, EEG, 2,'gui','off');
%Bandpass filter 45 hz
EEG = pop_eegfiltnew(EEG, [], 45, 38, 0, [], 1);
[ALLEEG EEG CURRENTSET] = pop_newset(ALLEEG, EEG, 3,'gui','off');
EEG = eeg_checkset( EEG );
%Extracting epochs and saving dataset
EEG = pop_epoch( EEG, {  }, [-1  2], 'newname', 'EDF file epochs',
'epochinfo', 'yes');
[ALLEEG EEG CURRENTSET] = pop_newset(ALLEEG, EEG,
4,'setname','before_artifacting','savenew','C:\\Users\\Admin\\Dropbox\\Thes
is\\Data\\Eeglab projects\\ID26-02before_artifacting.set','gui','off');
EEG = eeg_checkset( EEG );
```

Calculation of spectral powers

```
%This script allows to computer spectral absolute power for an epoch. It
%first calculates absolute spectral power for 1st second offset(local
% baseline), 1st second onset and 2sd second onset.

%Once calculated, it computes the resulting oscillation absolute powers(((1
onset + 2 onset)/2)-1 offset)
%asymmetry is calculated on the resulting oscillation by subtracting right
%electrode powers to left electrode powers

%Resulting matrixes(resulting oscillation and asymmetry) are stored in
vectors(concatenation of matrix rows) for building a dataset.
```

```
%to change epoch, change third variable for the desired epoch in
EEG.data(m,(1:128),1)


% Defining matrix for 1 offset
nr=14;  % rows(electrodes)
nc=5;  % columns(frequency bands)
table0=zeros(nr,nc);   % preallocate
e= 2%number of epoch
%computing absolute power for each band of interest for each electrode and
%storing them in the matrix
for m =1 : 14 %number de electrodos
[spectra,freqs] = spectopo(EEG.data(m,(1:128),e),0,128);
        % delta=1-4, theta=4-8, alpha=8-13, beta=13-30, gamma=30-45
        deltaIdx  = find(freqs>1 & freqs<4);
        thetaIdx = find(freqs>4 & freqs<8);
        alphaIdx = find(freqs>8 & freqs<13);
        betaIdx  = find(freqs>13 & freqs<30);
        gammaIdx = find(freqs>30 & freqs<45);




        % compute absolute power
        deltaPower = 10^(mean(spectra(deltaIdx))/10);
        thetaPower = 10^(mean(spectra(thetaIdx))/10);
        alphaPower = 10^(mean(spectra(alphaIdx))/10);
        betaPower  = 10^(mean(spectra(betaIdx))/10);
        gammaPower = 10^(mean(spectra(gammaIdx))/10);

        % fill Result Matrix
table0(m,:)=[deltaPower, thetaPower, alphaPower, betaPower, gammaPower];

end

% Defining matrix for 1 onset
nr=14;  % rows(electrodes)
nc=5;  % columns(frequency bands)
table1=zeros(nr,nc);   % preallocate

for m =1 : 14 %numero de electrodos
[spectra,freqs] = spectopo(EEG.data(m,(129:256),e),0,128);
        % delta=1-4, theta=4-8, alpha=8-13, beta=13-30, gamma=30-45
        deltaIdx  = find(freqs>1 & freqs<4);
        thetaIdx = find(freqs>4 & freqs<8);
        alphaIdx = find(freqs>8 & freqs<13);
        betaIdx  = find(freqs>13 & freqs<30);
        gammaIdx = find(freqs>30 & freqs<45);




        % compute absolute power
        deltaPower = 10^(mean(spectra(deltaIdx))/10);
        thetaPower = 10^(mean(spectra(thetaIdx))/10);
        alphaPower = 10^(mean(spectra(alphaIdx))/10);
        betaPower  = 10^(mean(spectra(betaIdx))/10);
        gammaPower = 10^(mean(spectra(gammaIdx))/10);

        % fill Result Matrix
table1(m,:)=[deltaPower, thetaPower, alphaPower, betaPower, gammaPower];
```

```
end

% Defining matrix for 2 offset
nr=14;  % rows(electrodes)
nc=5;  % columns(frequency bands)
table2=zeros(nr,nc);   % preallocate

for m =1 : 14 %numero de electrodos
[spectra,freqs] = spectopo(EEG.data(m,(257:384),e),0,128);
        % delta=1-4, theta=4-8, alpha=8-13, beta=13-30, gamma=30-45
        deltaIdx  = find(freqs>1 & freqs<4);
        thetaIdx = find(freqs>4 & freqs<8);
        alphaIdx = find(freqs>8 & freqs<13);
        betaIdx  = find(freqs>13 & freqs<30);
        gammaIdx = find(freqs>30 & freqs<45);



        % compute absolute power
        deltaPower = 10^(mean(spectra(deltaIdx))/10);
        thetaPower = 10^(mean(spectra(thetaIdx))/10);
        alphaPower = 10^(mean(spectra(alphaIdx))/10);
        betaPower  = 10^(mean(spectra(betaIdx))/10);
        gammaPower = 10^(mean(spectra(gammaIdx))/10);

        % fill Result Matrix
table2(m,:)=[deltaPower, thetaPower, alphaPower, betaPower, gammaPower];

end


% Computing resulting oscillation
res_osc = ((table1 + table2)/2)- table0;

%Computing asymmetry resulting oscillation, Left side power minus right
%side powers

nr=7;  % rows,pairs of electrodes(AF3-AF4, ...)
nc=5;  % columns, frequency bands
asy_osc =zeros(nr,nc);
asy_osc(1,:) = [(res_osc(1,:)) - (res_osc(14,:))];
asy_osc(2,:) = [(res_osc(2,:)) - (res_osc(13,:))];
asy_osc(3,:) = [(res_osc(3,:)) - (res_osc(12,:))];
asy_osc(4,:) = [(res_osc(4,:)) - (res_osc(11,:))];
asy_osc(5,:) = [(res_osc(5,:)) - (res_osc(10,:))];
asy_osc(6,:) = [(res_osc(6,:)) - (res_osc(9,:))];
asy_osc(7,:) = [(res_osc(7,:)) - (res_osc(8,:))];



%Creating feature vectors for Pattern recognition datasets
%Frequency_index
idx_fre= [res_osc(1,:), res_osc(2,:), res_osc(3,:),res_osc(4,:)...
    ,res_osc(5,:), res_osc(6,:),res_osc(7,:),res_osc(8,:)...
    ,res_osc(9,:), res_osc(10,:),
res_osc(11,:),res_osc(12,:),res_osc(13,:)...
    ,res_osc(14,:)];
```

```
%Asymmetry_index
idx_asy= [asy_osc(1,:), asy_osc(2,:), asy_osc(3,:),asy_osc(4,:)...
    ,asy_osc(5,:), asy_osc(6,:),asy_osc(7,:)];
```

PRtools Code for the 4 emotion ASi dataset – Courtesy of Michael Boelstoft Holte

```
%Reading datasets

pathName='C:\Documents and Settings\Administrador\Mis documentos\MATLAB\';
fid = fopen('4emotion_ASi.txt');
X =
textscan(fid,'%s%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%f%
f%f%f%f%f%f');
fclose(fid);
labels = X{1};
data =
[X{2},X{3},X{4},X{5},X{6},X{7},X{8},X{9},X{10},X{11},X{12},X{13},X{14},...

X{15},X{16},X{17},X{18},X{19},X{20},X{21},X{22},X{23},X{24},X{25},X{26},...
    X{27},X{28},X{29},X{30},X{31},X{32},X{33},X{34},X{35},X{36}];

z =dataset(data,labels);
prwaitbar off %it is bugged
%%
%Calculating accuracy from each of the classifiers

k = 10; %k-fold cross validation. k = 215 => leave-one-out cross validation
c_n={'nmc','ldc','qdc','knnc','svcp','svcr'}; %Classifiers names
for i = 1:length(c_n)
    if strcmp('nmc',c_n{i}),
        w = nmsc([]);
    elseif strcmp('ldc',c_n{i}),
        w = ldc([]);
    elseif strcmp('qdc',c_n{i}),
        w = qdc([]);
    elseif strcmp('knnc',c_n{i}),
        w = knnc([]);
    elseif strcmp('svcp',c_n{i}),
        w = nusvc([],proxm([],'p',1));
    elseif strcmp('svcr',c_n{i}),
        w = rbsvc([]);
    end

    [e(i),ce(:,i),nlab_out(:,i)] = crossval(z,w,k); %Crosvalidation
    acc(i) = 1-e(i) %Computing accuracy
    confmat(z.nlab, nlab_out(:,i)) %Confusion matrix

end
```

Bibliography

Adrian, E. D., & Matthews, B. H. C. (1934). The interpretation of potential waves in the cortex. *The Journal of physiology*, *81*(4), 440.

Bernhaupt, R. (2015). User Experience Evaluation Methods in the Games Development Life Cycle. In *Game User Experience Evaluation* (pp. 1-8). Springer International Publishing.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, *25*(1), 49-59.

Bradley, M., & Lang, P. J. (1999). *The International affective digitized sounds (IADS)[: stimuli, instruction manual and affective ratings*. NIMH Center for the Study of Emotion and Attention.

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, *45*(4), 624-634.

Brown, L., Grundlehner, B., & Penders, J. (2011, August). Towards wireless emotional valence detection from EEG. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE* (pp. 2188-2191). IEEE.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (Eds.). (2007). *Handbook of psychophysiology*. Cambridge University Press.

Chatrian, G. E., Lettich, E., & Nelson, P. L. (1985). Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities.*American Journal of EEG technology*, *25*(2), 83-92.

Coan, J. A., & Allen, J. J. (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. *Biological psychology*, *67*(1), 7-50.

Cohen, X. M (2014). Analysing neural time series data: Theory and Practice. Massachusetts. he MIT Press.

Davidson, R. J. (1992). Anterior cerebral asymmetry and the nature of emotion.*Brain and cognition*, *20*(1), 125-151.

Davidson, R. J. (2004). What does the prefrontal cortex “do” in affect: perspectives on frontal EEG asymmetry research. *Biological psychology*,*67*(1), 219-234.

Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology: I. *Journal of personality and social psychology*, *58*(2), 330.

Drachen, A. (2015). Behavioral Telemetry in Games User Research. In *Game User Experience Evaluation* (pp. 135-165). Springer International Publishing.

Drachen, A., El-Nasr, M. S., & Canossa, A. (2013). Game analytics–the basics. In *Game Analytics* (pp. 13-40). Springer London.

Drachen, A., Nacke, L. E., Yannakakis, G., & Pedersen, A. L. (2010, July). Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games* (pp. 49-54). ACM.

Duin, R. P. W., Juszczak, P., de Ridder, D., Paclık, P., Pezkalska, E., & Tax, D. M. J. (2004). Pr-tools. *Pattern Recognition Tools. http://www.prtools.org.*

EA (1998). Need for Speed III: Hot Pursuit. Seattle: EA

Emotiv Xavier TestBench [Computer Software]. (2015). Emotiv

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., ... & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, *53*(4), 712.

Ekman, P. (1999). Basic Emotions In T. Dalgleish and T. Power (Eds.) The Handbook of Cognition and Emotion Pp. 45-60.

Entertainment Software Association. (2015, April 15). Essential facts About the Computer and Video Game Industry. Retrieved from http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf

Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with computers*, *21*(1), 133-145

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral cortex*, *7*(4), 374-385.

Haas, L. F. (2003). Hans Berger (1873–1941), Richard Caton (1842–1926), and electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry*,*74*(1), 9-9.

Hadjidimitriou, S. K., & Hadjileontiadis, L. J. (2012). Toward an EEG-based recognition of music liking using time-frequency analysis. *Biomedical Engineering, IEEE Transactions on*, *59*(12), 3498-3510, L. J. (2012). Toward an EEG-based recognition of music liking using time-frequency analysis. *Biomedical Engineering, IEEE Transactions on*, *59*(12), 3498-3510

Hazlett, R. L. (2006, April). Measuring emotional valence during interactive experiences: boys at video game play. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 1023-1026). ACM.

He, E. J., Yuan, H., Yang, L., Sheikholeslami, C., & He, B. (2008, May). EEG spatio-spectral mapping during video game play. In *Information Technology and Applications in Biomedicine, 2008. ITAB 2008. International Conference on* (pp. 346-348). IEEE.

Hochleitner, C., Hochleitner, W., Graf, C., & Tscheligi, M. (2015). A Heuristic Framework for Evaluating User Experience in Games. In *Game User Experience Evaluation* (pp. 187-206). Springer International Publishing.

Jasper, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and clinical neurophysiology*, *10*, 371-375.

Järvinen, A. (2008). *Games without frontiers: Theories and methods for game studies and design*. Tampere University Press.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., Mckeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*(02), 163-178.

Kim, M. K., Kim, M., Oh, E., & Kim, S. P. (2013). A review on the computational methods for emotional state estimation from the human EEG.*Computational and mathematical methods in medicine*, *2013*.

Kitzinger, J. (1995). Qualitative research. Introducing focus groups. *BMJ: British medical journal*, *311*(7000), 299.

Koelstra, S., Yazdani, A., Soleymani, M., Mühl, C., Lee, J. S., Nijholt, A., ... & Patras, I. (2010). Single trial classification of EEG and peripheral physiological signals for recognition of emotions induced by music videos. In *Brain informatics* (pp. 89-100). Springer Berlin Heidelberg.

Konar, A., Halder, A., & Chakraborty, A. (2014). Introduction to Emotion Recognition. *Emotion Recognition: A Pattern Analysis Approach*, 1-45.

Konstantinidis, E. I., Frantzidis, C. A., Pappas, C., & Bamidis, P. D. (2012). Real time emotion aware applications: a case study employing emotion evocative pictures and neuro-physiological sensing enhanced by Graphic Processor Units. *Computer methods and programs in biomedicine*, *107*(1), 16-27.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual.*Technical report A-8*.

Lewis, C., & Mack, R. (1982, March). Learning to use a text processing system: Evidence from "thinking aloud" protocols. In *Proceedings of the 1982 conference on Human factors in computing systems* (pp. 387-392). ACM.

Lang, Peter J. "The emotion probe: studies of motivation and attention."*American psychologist* 50.5 (1995): 372.

Lin, Y. P., Wang, C. H., Jung, T. P., Wu, T. L., Jeng, S. K., Duann, J. R., & Chen, J. H. (2010). EEG-based emotion recognition in music listening.*Biomedical Engineering, IEEE Transactions on*, *57*(7), 1798-1806.

Makeig, S., Bell, A. J., Jung, T. P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. *Advances in neural information processing systems*, 145-151.

Malik, A. S., Osman, D. A., Pauzi, A. A., & Khairuddin, R. H. R. (2012, June). Investigating brain activation with respect to playing video games on large screens. In *Intelligent and Advanced Systems (ICIAS), 2012 4th International Conference on* (Vol. 1, pp. 86-90). IEEE.

Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & information technology*, *25*(2), 141-158.

Matlab [Computer Software]. (2015). The Mathworks

McAllister, G., & White, G. R. (2010). Video game development and user experience. In *Evaluating User Experience in Games* (pp. 107-128). Springer London.

Mirza-Babaei, P., Long, S., Foley, E., & McAllister, G. (2011, September). Understanding the contribution of biometrics to games user research. In *Proc. DiGRA* (pp. 329-347).

Murugappan, M., Nagarajan, R., & Yaacob, S. (2011). Combining spatial filtering and wavelet transform for classifying human emotions using EEG Signals. *Journal of Medical and Biological Engineering*, *31*(1), 45-51.

Nacke, L. E. (2015). Games User Research and Physiological Game Evaluation. In *Game User Experience Evaluation* (pp. 63-86). Springer International Publishing.

Nacke, L. E. (2013). An introduction to physiological player metrics for evaluating games. In *Game Analytics* (pp. 585-619). Springer London.

Nacke, L. E. (2010, May). Wiimote vs. controller: electroencephalographic measurement of affective gameplay interaction. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology* (pp. 159-166). ACM.

Nacke, L. E., Stellmach, S., & Lindley, C. A. (2010). Electroencephalographic assessment of player experience: A pilot study in affective ludology.*Simulation & Gaming*.

Nielsen, J. (1994, April). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414). ACM.

Nunez, P. L., Srinivasan, R., Westdorp, A. F., Wijesinghe, R. S., Tucker, D. M., Silberstein, R. B., & Cadusch, P. J. (1997). EEG coherency: I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalography and clinical neurophysiology*, *103*(5), 499-515.

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical neurophysiology*, *112*(4), 713-719.

Park, K. S., Choi, H., Lee, K. J., Lee, J. Y., An, K. O., & Kim, E. J. (2011). Emotion recognition based on the asymmetric left and right activation.*International Journal of Medicine and Medical Sciences*, *3*(6), 201-209.

Petrantonakis, P. C., & Hadjileontiadis, L. J. (2011). A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition. *Information Technology in Biomedicine, IEEE Transactions on*,*15*(5), 737-746.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Theories of emotion*, *1*.

Preece, J., Sharp, H., & Rogers, Y. (2015). *Interaction Design-beyond human-computer interaction*. John Wiley & Sons.

Ravaja, N., Saari, T., Salminen, M., Laarni, J., & Kallinen, K. (2006). Phasic emotional reactions to video game events: A psychophysiological investigation.*Media Psychology*, *8*(4), 343-367.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169-192.

Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of research in Personality*, *11*(3), 273-294.

Salminen, M., & Ravaja, N. (2007). Oscillatory brain responses evoked by video game events: The case of Super Monkey Ball 2. *CyberPsychology & Behavior*, *10*(3), 330-338.

Salminen, M., & Ravaja, N. (2008). Increased oscillatory theta activation evoked by violent digital game events. *Neuroscience letters*, *435*(1), 69-72.

Schwartz, B. J., Salustri, C., Kaufman, L., & Williamson, S. J. (1989). Alpha suppression related to a cognitive task. In *Advances in Biomagnetism* (pp. 237-240). Springer US.

Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, *3*(3), 3-3.

Schacter, D. L. (1977). EEG theta waves and psychological phenomena: A review and analysis. *Biological psychology*, *5*(1), 47-82.

Schier, M. A. (2000). Changes in EEG alpha power during simulated driving: a demonstration. *International Journal of Psychophysiology*, *37*(2), 155-162.

Takatalo, J., Häkkinen, J., & Nyman, G. (2015). Understanding Presence, Involvement, and Flow in Digital Games. In *Game User Experience Evaluation*(pp. 87-111). Springer International Publishing.

Team17 (2012). Worms Revolution. Wakefield:Team17

Van Der Heijden, F., Duin, R., De Ridder, D., & Tax, D. M. (2005).*Classification, parameter estimation and state estimation: an engineering approach using MATLAB*. John Wiley & Sons.

Wang, X. W., Nie, D., & Lu, B. L. (2011, January). EEG-based emotion recognition using frequency domain features and support vector machines. In*Neural Information Processing* (pp. 734-743). Springer Berlin Heidelberg.

Wallner, G., & Kriglstein, S. (2015, January). An introduction to gameplay data visualization. In *Game Research Methods* (pp. 231-250). ETC Press.

Wallner, G., & Kriglstein, S. (2013). Visualization-based analysis of gameplay data–a review of literature. *Entertainment Computing*, *4*(3), 143-155.

Wehbe, R. R., & Nacke, L. (2013). An Introduction to EEG Analysis Techniques and Brain-Computer Interfaces for Games User Researchers.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments*, *7*(3), 225-240.

What Is BCI and How Did It Evolve? (2015, June 19).Retrieved from http://neurosky.com/2015/06/what-is-bci-and-how-did-it-evolve/