THE DEVELOPMENT OF A MULTIMODAL NEUROADAPTIVE GAMING TECHNOLOGY TO DISTRACT FROM PAINFUL EXPERIENCES.

Kellyann Stamp BSc.

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Doctor of Philosophy

April 2020

Table of Contents

List of Figures

List of Tables

Abstract

Painful experiences can be mitigated by distraction techniques such as video game distraction, due to limited available attentional resources. There are many benefits to using video games as a non-pharmacological intervention, including their cost-effectiveness and absence of side effects or withdrawal symptoms. However, video games cannot provide a distraction which is sufficient for pain management if they are not engaging. This work aims to discuss how and why video games capture attention and explore how modulating game factors can affect the response to pain. The aim of this work in its entirety is to develop a neuroadaptive game which is tailored to reorient attention away from a painful experience, and towards the distraction technique. The neuroadaptive element of this technology will enable a balance of challenge and skill which make a unique and playable game for each participant. The development of the neuroadaptive game was supported by two studies. Study One focused on the determination of optimal game difficulty level for pain distraction, and Study Two furthered this research, alongside determining optimal neurological sites for the monitoring of attention and attentional reorientation. Study 3 explored the use of a neuroadaptive gaming technology to distract from pain – a bespoke, real-time data processing pipeline was developed for this purpose. The limitations of the neuroadaptive game are discussed in detail with considerations for future work and development. The results of the three studies carried out during the course of this work indicate that real-time pre-processing and classification of fNIRS data to a good standard is possible. The studies also revealed that the montage for data collection and features used for data collection are crucial considerations for classification accuracy. This thesis also has implications for further work into neuroadaptive technologies and how these systems can be tested and verified. Statistical significance between a non-neuroadaptive game and a neuroadaptive game was not found throughout the course of this work, although the potential explanations and future considerations are discussed in detail. Overall, we were able to confirm that pain tolerance can be improved with the use of a distraction task, but that the balance of task difficulty and skill level is delicate and requires further exploration.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Acknowledgements

Firstly, I would like to thank my supervisory team, Dr Chelsea Dobbins, Professor Stephen Fairclough and Dr Helen Poole, without whom this research would not have been possible. Their expertise, support, understanding and encouragement has enabled me to complete this research to the best of my ability. The input of my supervisory team has proven invaluable and I know that their knowledge and advice will support me throughout my career. I would especially like to thank Dr Chelsea Dobbins for continuing to provide her invaluable support during and after moving halfway across the world – I really appreciate that you still found the time to support me on this journey. I would also like to acknowledge my appreciation for Professor Stephen Fairclough for taking on the extra work of becoming my Director of Studies part-way through this journey – the support, encouragement and understanding made the transition smooth and ensured that I was always supported throughout this journey. I would like to thank Dr Helen Poole for always finding the time to support my research and provide encouragement and support. Without my supervisory team, this Thesis would not have been possible, and I certainly wouldn't have enjoyed the journey anywhere near as much. I am lucky to have found such a supportive and knowledgeable team.

I would also like to thank all of my family and friends for their support throughout the last few years. Special thanks to my parents for letting me stay at home and eat all of their food. Without the support, understanding and encouragement of my parents, I would never have dreamed of embarking on this journey in the first place. Thank you especially to my dad for teaching me how to use a computer (and sparking an interest that would lead me to the completion of this PhD), and to my mum for putting up with the house full of computers that came along with that interest! Special thanks also to my brothers, Robbie & Tommy, for their support and love. To Mike, Abbie & Kate – I would like to say thank you for understanding my strange schedule and boring conversation topics – and for only asking me a few times when I would finally finish my Thesis!

Finally, I would like to thank the entire Department of Computer Science & School of Psychology for supporting me throughout this journey, for their advice and support, and for their willingness to take part in my studies! I would like to thank Thomas Berry for encouraging me to apply for PhD study and supporting me throughout the journey, and Chris Baker & Rachel Hagan for their unending support and encouragement.

Publications Resulting from This Thesis

- Stephen Fairclough, **Kellyann Stamp,** Chelsea Dobbins, Helen Poole "Computer Games as Distraction from Pain: Effects of Hardware and Difficulty on Pain Tolerance and Subjective Immersion" in *International Journal of Human Computer Studies,* vol. 139, p. 102427, July 2020

- **Kellyann Stamp**, Chelsea Dobbins, Stephen Fairclough "Utilization of Neurophysiological Data to Classify Player Immersion to Distraction from Pain" in *Human Computer Interaction International (HCII'20)*, 2020 (Accepted)

- Stephen Fairclough, **Kellyann Stamp**, Chelsea Dobbins "Neuroadaptive Gaming in the Context of Pain" in *Human Computer Interaction International (HCII'20)*, 2020 (Accepted)

- **Kellyann Stamp**, Chelsea Dobbins, Stephen Fairclough and Helen Poole "The Influence of Game Demand on Distraction from Experimental Pain: A fNIRS Study" in *Frontiers in Human Neuroscience Conference Abstract: 2nd International Neuroergonomics Conference*, 2018

- **Kellyann Stamp**, Chelsea Dobbins, Stephen Fairclough and Helen Poole "Development of A Neuroadaptive Gaming Technology to Distract from Painful Procedures" in *1st Neuroadaptive Technology Conference 2017 (NAT'17),* 2017

Chapter 1 - Introduction

1.1 The Pain Experience and Theory of Distraction

The experience of pain can have long lasting and detrimental effects on the sufferer. This is particularly true during the medical treatment of children, who can be affected for the rest of their lives by painful experiences they may encounter during childhood (Broad & Wheeler, 2006; Diseth, 2006). Opiate administration is a common pain management method used during medical procedures (Cooper et al., 2017; Matson, Johnson, Tran, Horton, & Sterner-Allison, 2019). Although it is crucial that pain is adequately managed during medical interventions, the use of opioids for pain management is not without drawbacks. One common side effect of repeated opioid treatment is a growing tolerance to opioid medication, wherein the day 1 effects of opioid administration are more potent than administration in the following days. If a painful treatment must be repeated multiple times (for example, during the changing of dressings for a burns patient) then a higher dose of opiate medication may eventually be required to achieve the desired analgesic effect. Tolerance comes hand-in-hand with the potential for experiencing withdrawal when medication is reduced or removed. The effects of opioid withdrawal are not only physical (including tremors, grimacing, vomiting and other gastrointestinal distress) but can also be mental (agitation, anxiety and insomnia.) Although it is obviously important that pain is managed to prevent further negative effects, the negative effects of some types of pain management cannot be ignored (Anand et al., 2010). Therefore, it is important to consider alternative or complimentary interventions, which could both reduce the experience of pain and mitigate the use of opioids.

Distraction is one way in which pain can be mitigated during painful clinical procedures. Tasks designed to distract from pain, such as playing with toys or blowing bubbles, are commonly used and considered to be clinically effective in distracting from pain (Goldman & Lloyd-Thomas, 2004; Gupta et al., 2006; Williams & Ishimine, 2016). Distraction is also anecdotally used to prevent children from attending to non-serious injuries, e.g. parents pulling focus away from an accident by changing the subject or introducing a new stimulus (Goldman & Lloyd-Thomas, 2004). The utility of distraction as a method for pain management is based upon an assumption that selective attention plays a role in pain perception and can be modulated by distracting activities. Attention is regarded as a limited resource and so it is difficult to attend to multiple tasks that are presented simultaneously (Eccleston & Crombez, 1999; Scalf, Torralbo, Tapia, & Beck, 2013). If a person is focused on one task, then external stimuli, which have no relevance to the goal associated with that task, will be filtered to maintain task focus. The reduction of attentional processing on non-task related stimuli means that all of the available attention can be focused on the primary goal of the task (Eccleston & Crombez, 1999). Norman and Shalice (1986) (Norman & Shallice, 1980) proposed that the selection of where to direct attentional resources (i.e. which stimuli) is an automatically triggered behaviour, which is only interrupted by either the: completion of a goal, inability to reach a goal, or a shift to secondary stimuli that are more immediately relevant than the current goal. The level of attention focused on the primary goal has a direct effect on the salience of secondary stimuli. For example, the phenomena of inattention blindness and deafness have been observed in pilots during flight under difficult conditions. This is due to the conditions of the flight task being demanding. It has been observed that pilots may inadvertently ignore auditory warnings, which signal further action is required (Durantin, Dehais, Gonthier, Terzibas, & Callan, 2017). Inattentional deafness is an example of secondary stimuli not being perceived due to the high level of attention focused on primary stimuli. The phenomena of intentional blindness and deafness reinforces the theory that attentional capacity is limited. Attending to multiple stimuli becomes more difficult when the primary stimuli presents a high level of demand. Although inattentional deafness (and blindness) pose issues when a person is required to carry out a task that may contain various forms of stimuli which are all competing for attention, there are benefits to this attentional control system. Broadbent's (1956) (Broadbent, 1956) theory of an attentional bottleneck describes how goal-irrelevant stimuli can remain unprocessed and therefore unobserved due to their irrelevant nature. This theory is useful when considered alongside painful experiences. The theory of the attentional bottleneck implies that, if a person is entirely focused on achieving a goal, they will not process irrelevant information. This, in turn, implies that a person may devote less attention to a source of painful sensation if they are entirely focused on a different primary task.

Pain has an interruptive function that demands attention, even in the presence of primary tasks (Legrain et al., 2009). The experience of pain evokes a primitive desire to escape from the painful stimuli. In addition, pain has the ability to override current attentional goals and to promote the goal to remove oneself from a painful situation (Crombez, Van Ryckeghem, Eccleston, & Van Damme, 2013). The capacity of painful stimulation to capture attention depends on various characteristics of the painful experience, such as the novelty and intensity of pain (Keogh, Moore, Duggan, Payne, & Eccleston, 2013). Pain can affect both top-down and bottom-up modes of selective attention, which refers to how the brain processes stimuli that demands attention (Theeuwes, 1991). Top-down control relates to a current task and allows a person to focus on stimuli that relates directly to the task, rather than alternate stimuli that may also be present. Bottom-up control relates to the ability of non-task relevant stimuli to attract and capture the attention of an individual (Legrain et al., 2009). Top-down control can often be seen as voluntary, as a person has to make a decision to focus on a task, whereas bottom-up control is often involuntary and relates to the re-orientation of attention from a task towards salient external stimuli. However, the salience of the secondary stimuli must be sufficiently important to initiate the process of re-orientation (Tiemann et al., 2015). To relate theories of attentional control back to previously discussed research (Dehais et al., 2014), a pilot focusing on a landing task exerts a high level of focus (top-down control) on the task, and therefore attention is not re-directed (via bottom up control) to the external stimulus of an alarm bell (causing inattentional deafness). In this case, attention is not re-directed because the new stimulus is not sufficiently salient to interrupt the primary task. However, pain is an extremely salient experience, which means that pain does have the ability to re-direct attention (via bottom up control) that can, in turn, modify the current focus away from the task goal, and towards the painful experience. In this case, top-down control causes the individual to cease concentrating on their original task and switch attention to how they can terminate or reduce the painful experience. Nevertheless, it is already understood that attention is a limited resource. Therefore, a task that is inherently demanding for an individual engages the top-down control mechanism, which supresses activity from the bottom-up control mechanism in order to reduce the potential of interruptions.

The demand of a task is often referred to as *task load*, which has been studied previously to examine the interactions between task demand and the observance of task-unrelated stimuli. The relevance of task demand, in regard to the ability of the task to distract from a painful experience, has been extensively reviewed. Findings have indicated that high task load is required before non-task related stimuli can be subconsciously ignored by an individual (Lavie, 2005; Lavie, Beck, & Konstantinou, 2014). In order to create a distraction that could prevent a person from observing a painful experience, the task must demand a high level of load from the participant. The level of load that a task demands can be modulated via a variety of factors.

Distraction techniques can typically be split into two categories – *active* and *passive* (MacLaren & Cohen, 2005; Nilsson, Enskär, Hallqvist, & Kokinsky, 2013). Passive distraction techniques refer to techniques that do not require active input from a person. An example would be watching the television. Regardless of the level of attention that a person is exhibiting, the television will continue to play. An active distraction refers to a task that requires the active engagement of the person. For example, if a person is reading a book but then becomes disengaged or distracted by another stimulus, then the book is no longer being read. Studies have indicated that active distractions are far more successful at distracting people from pain than passive distractions (Hussein, 2015; Wohlheiter & Dahlquist, 2013). Active engagement with a task requires higher activation of top-down control, in comparison to passive tasks, thus leaving a reduced amount of attention available to attend to an activity from the bottom-up control network.

In order to sustain engagement towards a task that actively distracts from pain, intrinsic motivation is of the upmost importance. Intrinsic motivation (Adams, Little, & Ryan, 2017) refers to an individual's desire to complete a task not because they have been told that they have to, but because they themselves want to. Skill level is correlated with intrinsic motivation as an individual must possess the sufficient skill level to complete the task that they wish to undertake, in order for them to remain motivated (Rheinberg, 2008). For example, if a child is presented with a book that is too advanced for their literacy level, then they become frustrated because they cannot engage with the book. This does not mean, however, that they could not become engaged by reading another book - but it does indicate that there must be a linear relationship between the literacy level of the reader and the book. This observation supports a theory that the maximum level of distraction can be achieved if the task provides an optimal level of challenge. However, determining the optimal level of challenge requires a direct match between individual skill level and the demand of a task (Akizuki & Ohashi, 2015; Guadagnoll & Lee, 2004). Although this can be a difficult task, it is not without its benefits. A recent study (Meng, Pei, Zheng, & Ma, 2016) reported that an optimal challenge condition produced Electroencephalogram (EEG) results that indicated more concentrated attention and stronger intrinsic motivation, in comparison to a non-optimal challenge condition. These results were observed in the F6 and FC6, using the 10-20 representation of the Right Ventral Frontal Cortex.

Video games as a method of pain distraction have been extensively researched (**collate all refs here**). A 2011 study (Jameson, Trevena, & Swain, 2011a) found that the use of a video game increased participants pain tolerance (as measured by CPT submersion time), and lead to a reduction in reported pain. Participants also reported that they would be more willing to complete the CPT again if an active distraction was used, in comparison to the offer of an alternative (passive) distraction. Although this research is promising, it is important to note that this study only compared the effects of a video game (active task) to the effects of a passive task (watching TV) – therefore we cannot assume from this research that video games are more effective than any other active task. A 2016 study (Parker et al., 2016) found that the use of a video game did improve participants self-reported pain scores. This study focused on individuals who required rehabilitation due to acute burns. Participants were spilt into two groups – one group received standard rehabilitation exercise therapy, whereas the second group received this standard therapy alongside 5 days of twice-daily exercise using the Nintendo Wii. Overall, the group that received the Wii therapy reported a 17% drop between their pre and post therapy pain reporting compared to the group that only received standard therapy. However, it could be argued that the additional exercise time provided by the Wii therapy is at least partially responsible for this reduction in pain. Although this study also provides promising results, it is important to note that this was a long-term study and did not focus on the effects of distraction on pain response – but rather the long-term effects of a gaming intervention.

A 2016 study (Kaheni, Rezai, Bagheri-nesami, & Hossein, n.d.) found that children who experience a video game intervention when receiving burn dressing changes reported less pain than children in a control group who didn't receive any distraction intervention. Although these results are promising in terms of the effectiveness of a game distraction, they still do not answer the question as to weather a game distraction is more effective than an alternate type of distraction. A 2019 study (Inan & Inal, 2019) does provide more insight into the effects of different active distractions on children's response to pain. This study consisted of four groups of children (one group which played video games, one group which watched TV, one group which were distracted by their parents, and one group without distraction intervention). This study found that the lowest level of anxiety and pain perception was found in the Video Game group, and that both the TV group and the parental distraction group had lower pain and anxiety scores than the control group (no distraction). This study provides further evidence supporting both the efficacy of a distraction task, and the superiority of a video game distraction (in comparison to the parent-lead conversational distraction).

In 2011, a study (Gordon, Merchant, Zanbaka, Hodges, & Goolkasian, 2011) was carried out to determine how the presentation of a video game may affect its effect on pain tolerance. This study used the same video game (presented in VR and projected onto a wall) and compared the differences in display type, as well as differences in comparison to a positive emotion induction condition (non-game intervention). This study found that both the VR game and the projected game significantly reduced self-reported pain scores – however, there was no significant difference between the way this game was displayed. Further research into this topic (S. H. Fairclough, Stamp, Dobbins, & Poole, 2020) was carried out which found similar results – the display and audio settings of a game did not significantly affect pain tolerance. However, this study did indicate that the level of game difficulty did significantly affect the response to pain. If a game provided a higher level of challenge, then participants experienced a significant increase in pain tolerance. This research indicates tha the experience of challenge-based immersion may be the most effective way to increase a participant's pain tolerance.

Studies have shown that video games are the most effective form of active distraction (Weiss, Dahlquist, & Wohlheiter, 2011) due to the experience of immersion. Video games result in larger increases in pain tolerance and decreases in pain perception than tasks such as watching television (Jameson, Trevena, & Swain, 2011b). Therefore, it can be assumed that the most effective form of active distraction would be an adaptive video game. We assume that an adaptive video game would be further effective due to the tailored nature of the gameplay experience, meaning that skill level is not a barrier for attainment. However, designing an adaptive game to distract from pain requires a variety of different considerations. Firstly, it must be understood how a video game captures attention, and secondly, it must be understood how a high level of attention can be maintained via game play. Statistics suggest that gamers spend an average of around 7 hours a week playing video games [1], whilst Przybylski et al. (2010) (Przybylski, Deci, Rigby, & Ryan, 2014) suggest that many people spend so much time playing video games because they fulfil basic psychological needs for the players. The needs that video games are perceived to satisfy are defined by the Self Determination Theory (SDT) (Adams et al., 2017), which includes: Competence, Autonomy and Mastery:

- *Competency* refers to feeling capable. Even the earliest video games were designed specifically to make the player feel able, by gradually increasing the level of difficulty of the game to ensure that players did not feel overwhelmed by the demands of the game, but still felt as though they were exerting effort towards the game. Competency is still a relevant factor of video game design today, with techniques such as player matching being utilised to ensure that players compete with players with a similar skill level.

- *Autonomy* refers to the feeling of choice within the game environment. As video game design has advanced, one major focus that has remained relevant has been the experience of choice. As time has passed, video game environments have become larger and have incorporated more choice-related options. Many video games now give the player the opportunity to explore a variety of different outcomes to both in-game tasks and the overall game depending on the choices that they make (Somerdin 2016). Many games are now designed to appear limitless to a player, which supports the feelings of autonomy that human beings crave.

- *Relatedness* refers to the sense that, rather than being an individual, the player is part of a larger group who are working together to reach the same goal. This sense of relatedness is achieved via the creation of social bonds, whether these are from simple text-based outputs, or more advanced technologies, such as Voice over Internet Protocol (VoIP).

[1] https://www.limelight.com/resources/white-paper/state-of-online-gaming-2019/

Games that satisfy players SDT needs are more likely to encourage a player's intrinsic motivation, which is an important element required for creating a distraction task. When considering how a game can provide the most potent distraction for a player, it is important to consider *immersion*. Immersion is a model of graded attention (Jennett et al., 2008), which begins with the experience of engagement, before moving on to engrossment, and finally total immersion, which can also be referred to as flow. The desired state for maximum distraction is *total immersion*, where it would be expected that the phenomena of inattentional blindness and deafness to any non-game related stimuli would be observed. A player cannot simply become immersed in a game as soon as they begin playing, but must first experience the preceding two states before immersion occurs. A brief description of the three states of immersion is as follows (Cairns, Cox, & Imran Nordin, 2014)

- *Engagement* relates to the investment of time, effort and attention towards a task. The major barrier at this point, which may prevent a player from entering a state of engagement, is access. In this case, access refers both to a player's preference in game style and access to the controls of a game; if a player cannot master the physical control of a game, they will also be unable to engage.

- *Engrossment* refers to the feeling of being enthralled with a game. In order for a person to become engrossed with a game, they must begin to be emotionally affected by the game. This does not necessarily mean that the player feels an emotional attachment towards characters of the game, but rather that they gain respect for the game and the effort that has been applied to its creation. Good game construction is critical for achieving engrossment. If a player can experience engrossment, they will start to lose their sense of self awareness and situational awareness. The player is still aware that they are playing a game and are not within the game, but real-world stimuli become less relevant.

- *Total Immersion* is often referred to as presence and is the final stage to achieving the experience of immersion. Total immersion has been described as the feeling of no longer existing in the real world, but rather existing within the game or the computer. In order to achieve immersion, a player must experience empathy towards a character or team and experience the emotions that the character is presumed to be experiencing.

The theory of total immersion is supported by the flow theory, which was described by Csikszentmihalyi (1975) (Csikszentmihalyi, 1975). Flow theory proposes that the goals of future achievement may be less important than a sense of fulfilment, which is associated with competence and autonomy, in order to motivate a person to continue a task. For example, a person has not entered the flow state just because they are immersed, but rather because their immersion stems from their current enjoyment of the game, and not their desire to reach a goal at some future point. For flow to be achieved, there must be an intrinsic enjoyment in the entire task, rather than just a desire to succeed. The major barrier that may prevent a player from experiencing or entering the flow state is the balance between challenge and skills (Nacke & Lindley, 2008). Figure 1 below illustrates this barrier: this figure visualises how the level of skill and level of game difficulty are correlated, wherein low skill and high game difficulty encourage anxiety, high skill and low difficulty encourage boredom, and an evenly balanced level of skill and game difficulty can encourage a player to enter the flow state.

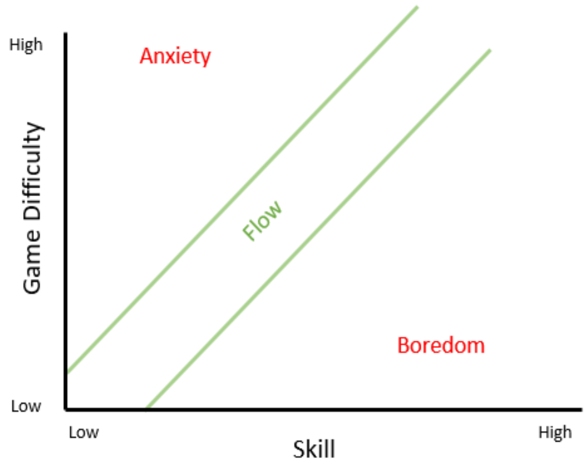


Figure 1 - Flow Model (adapted from Nacke and Lindley 2008)

This section has discussed how attention is important for pain perception. An individual only has a limited amount of attention available to expend on stimuli, which means that attention can be manipulated towards distraction tasks and away from the sensation of pain. Distractions can work as an analgesic by enhancing top-down control and task focus at the expense of bottom-up processes. Bottom up processes are associated with the interruptive function of pain, so reducing bottom-up control can reduce the likelihood of an individual attending to a painful experience. Tasks that require active engagement are more successful at enhancing top-down control than passive tasks, and tasks in which the level of challenge can be optimised (such as games) are further successful at enhancing this top-down control. Another benefit of games is that a good game performance can satisfy basic needs, such as competence, which further improves task engagement. From this analysis of gaming research, it is concluded that immersion (which is a graded description of top-down control) and flow are two important factors that can influence the efficacy of distraction. Whereas immersion focuses primarily on maximising top-down control, the theory of flow argues also that an intrinsic level of enjoyment is also necessary to create the type of substantial distraction associated with engrossment or total immersion. In addition, it is equally important that the player has a sufficient level of skill to meet the challenge of the game before distraction can be fully achieved.

1.2 Analgesic Technology

Video games have been used as a distraction from pain over multiple studies, using both clinical and experimental pain (Hoffman, Doctor, Patterson, Carrougher, & Furness, 2000; Law et al., 2011;

Miller, Rodger, Bucolo, Greer, & Kimble, 2010; Weiss et al., 2011). Past research has explored how technology can be used as an analgesic by distracting the person from pain. For example, severe burn injuries are amongst some of the most painful injuries that a person can experience. The healing process can include multiple weeks of frequent dressing changes and skin stretching. The pain experienced during these procedures is so intense that opioid use may be insufficient, even prior to the build-up of tolerance. In 1996, Hoffman et al. (Hoffman, 2004) conducted a study at Harborview Burn Centre in Seattle to determine if patients could experience pain alleviation via the use of virtual reality (VR) technology. The initial study was carried out with two teenage male participants who both had severe burns over their bodies. The study was conducted whilst the patients had staples removed from skin grafts. Each participant still received the standard level of opioid treatment prior to the procedure and also played two distraction games (one VR and one non-VR) during the procedure. The VR distraction used was *Spiderworld*, which had originally been designed to attempt to combat a fear of spiders. The standard video game distraction was either *Wave Race 64* or *Mario Kart 64* (both of which were racing games). Both of the patients rated their pain lower during the VR condition than the game condition. The effects of distraction can also be observed via neurophysiological signals. For instance, Hoffman et al (2004) (Hoffman, 2004) used functional Magnetic Resonance Imaging (fMRI) to determine whether there were noticeable changes in the brain activation when participants experienced thermal pain with and without distraction via VR intervention. They reported that specific areas of the brain, such as the insula, thalamus, primary and secondary somatosensory cortices and areas of the anterior cingulate cortex, which are areas known to be associated with the perception of pain, were highly activated when the participant was experiencing pain without the VR distraction. However, when participants experienced distraction via VR, they reported subjective pain as less severe.

A number of studies have been conducted to explore the use of other active forms of distractions as analgesic interventions. Sprenger et al (2012) (Sprenger et al., 2012) carried out a study in which they monitored the cervical spinal cord via fMRI whilst participants experienced experimental thermal pain. Participants in this study underwent two working memory distraction tasks – a 1-back test and a 2-back test. It was theorised that the 2-back test would significantly modulate the response to pain, as higher working memory load would be required. As expected, participants rated their pain as significantly higher during the 1-back task as compared to the 2-back task. A significant reduction in the blood oxygen level dependant (BOLD) response was observed during the 2-back task compared to the 1-back task, which indicates that the effect of the distraction task affected opioid neurotransmission and actively provided an analgesic experience to the participants, without the use of a medication intervention. To further test this theory, the study was repeated using an opioid antagonist, which are drugs that are used to reverse the effects of opiates. This study was double blind, with neither the participant nor the experimenter being informed as to whether the participant was receiving an opioid antagonist or a placebo. The results of this study showed that pain intensity ratings given by the participants during the placebo trial were the same as those given in the initial study (a significant reduction in pain observed during the 2-back task). However, following the administration of an opioid antagonist, the pain perception results of the 1-back task were similar to those in the original study, but pain perception was rated significantly higher during the 2-back task. These results further support the hypothesis that opioid neurotransmission is occurring when a distraction task is used and can therefore be blocked by opioid antagonists.

A number of studies have been undertaken to determine how different kinds of distractions influence the pain response. One study (Abdelmoniem & Mahmoud, 2015) explored the use of a passive distraction (listening to music), an active distraction (alternate up-and-down movement of the legs) and passive-active distraction (a combination of both the passive and active distractions). The study measured 90 children's perception of pain during venepuncture using the Sound, Eyes and Motor scale (Wright, Weinberger, Marti, & Plotzke, 1991), and the Wong-Baker Faces scale (Wong & Baker, 1988). This study found no significant difference between either of the three groups. However, considering the literature relating to the required effectiveness of an active distraction task (the experience of flow and immersion, for example), this is not a surprising result. Although moving one's legs up and down alternately is an active distraction, it does not meet any of the criteria of providing an engaging, engrossing or immersive distraction. Another study (Ebrahimpour, Sadeghi, Najafi, Iraj, & Shahrokhi, 2015) focused on children who received once daily insulin injections. This study, which used an active distraction (video game) and a control group, found that children in the distraction group had a significant reduction in behavioural distress compared to the control group over the course of one week. It is important to also consider the ability of a distraction to sustain an analgesic effect over repeated experiences. Novelty can increase the ability of a task to distract, however, if a patient or participant is to receive multiple painful procedures, then distraction and novelty can reduce with each experience. However, in an 8-week study conducted by Rutter et al (2009) (Rutter, Dahlquist, & Weiss, 2009), it was reported that the reduction in pain perception and increase in pain tolerance due to VR distraction that were observed in Week 1 were equivalent to those observed by Week 8. This indicates that novelty was not significantly affecting the distraction properties of the task.

There is both neurophysiological and subjective evidence to support the claim that technology has the capability to function as an analgesic. Further evidence suggests that analgesic technology functions by distracting participants, and therefore activating top-down attentional control. However, the challenge for technology to work as an analgesic is to create maximum levels of distraction during a painful experience and to sustain distraction over repeated episodes of pain.

1.3 Neuroadaptive Technology

It has been established that games can function as an analgesic by actively distracting attention away from a painful sensation. The effect of the distraction task is maximised when the level of challenge is matched to the skill of the individual player. By tuning game demand to match the players' skill level, a game can promote high performance states of attention, such as engrossment and flow. One way of maximising this distraction is to create games that work on a closed-loop model and are capable of adapting the level of game difficulty in real-time, to match the skill level of player and sustain a high level of challenge. This section will consider different approaches to the creation of closed-loop technologies to distract from pain.

1.3.1 Open, Closed and Biocybernetic Loops

Neuroadaptive technology works on the basis of closed-loop control (Wiener, 1961). In the past, the brain has been considered to be a *black box*, which data can be fed into and received from. Treating the brain as though it were a black box has led to an influx of neurophysiological research in which open loops have been used. In an open-loop system, the stimulus is arbitrary, and the result of the stimulus (recorded from the brain) is the result, whereby the resulting brain activity has no bearing on the overall study protocol. In an open-loop experiment, no changes are made to the study protocol or the stimulus, regardless of the brain activity (Zrenner, Belardinelli, Müller-Dahlhaus, & Ziemann, 2016). This is in complete contrast to how the brain actually works in a real-world situation. In a natural environment, the effect that the stimulus has on the brain affects the future choices or decisions that a person will make. For example, if a person touches a hot bowl, their brain will process the pain they experience from the heat, and the person will understand that they will experience pain again if they touch the bowl again. This will, therefore, prevent the person from touching the bowl. An open-loop system does not measure this consequence-based decision making, as an open-loop is used purely to examine the effect of the stimulus, and not to make changes to the stimulus based on the brain response.

A closed-loop system works in a more realistic way and involves the real-time detection of states (based on psychophysiology, neurophysiology, or a combination of the two) and a translation of those states into a real-time adaptive response that is both timely and intuitive from the perspective of the user. In a closed-loop system, the stimulus is delivered to a participant, and their resulting neurophysiological activity is measured. The neurophysiological data is analysed in order to determine how the participant's brain responded to the stimulus and, based on this data, changes can be made to the stimulus in order to make it more appropriate for the participant. Figure *2* provides a representation of a closed-loop system.

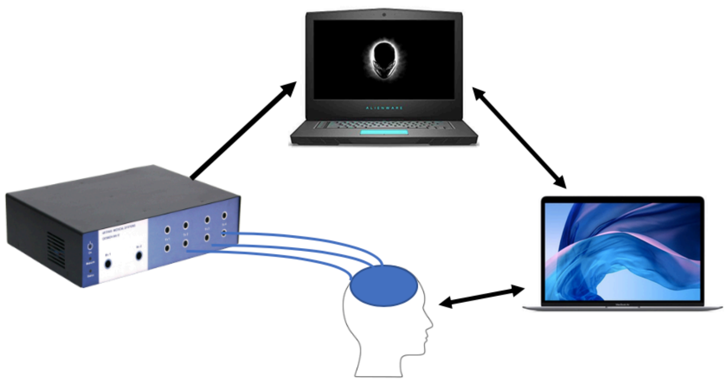


Figure 2 - Visualisation of a basic closed-loop system

When a closed-loop system is implemented in a neurophysiological system, it is also referred to as a biocybernetic loop. Biocybernetic loops are a facet of closed-loop systems that specifically relate to the use of human data (the 'bio' element) to make changes to automatically controlled systems and their communication (the 'cybernetic' element.) The biocybernetic loop was developed in 1994 by a team of scientists at NASA and Lockheed, led by Alan Pope (Pope, Bogart, & Bartolome, 1995). The purpose of creating the biocybernetic loop in this case was to make changes to an autopilot system. The biocybernetic loop was used to determine pilot engagement and adapt the autopilot functionality in order to maintain the engagement of the pilot. Although this study found that the biocybernetic loop was able to differentiate between different mental states, this study did not present any evidence that indicated that the system affected pilot performance. In 1999, Freeman et al. (Freeman, Mikulka, Prinzel, & Scerbo, 1999) replicated Pope et al.'s work (Pope et al., 1995) in order to establish whether the adaptive system would improve overall performance. Pope et al.'s 1995 study measured changes in EEG, via the slope of the signal, however, Freeman et al. (Freeman et al., 1999) argued that a system such as this may be too sensitive to small changes in brain activity and therefore cause incorrect adaptations. For this reason, Freeman et al (Freeman et al., 1999) suggested that the adaptive system should firstly record a baseline measure of brain activity, and then make adaptations based on significant deviations from this baseline. In this study, both positive and negative feedback loops were used. In BCI, a positive feedback loop is a loop which adjusts in one direction in order to achieve a pre-defined state – actions which have previously achieved a positive reaction will be repeated. A negative feedback loop assesses the current mental state and makes adjustments in either direction in order to attain and maintain a desired state. Both positive and negative feedback loops have potential negative implications – positive feedback loops can continue to make the same (previously effective) changes regardless of whether they remain affective; negative feedback loops rely on finding an acceptable range, which means that response times can be slow and therefore cause a BCI to become temporarily ineffective. The results of this study indicated that negative feedback loops were more effective than positive ones, and that there was an improvement found in overall performance when the adaptive system was used (Scerbo, Freeman, & Mikulka, 2003). Fairclough (S. H. Fairclough, 2009a) expanded on the concept of the biocybernetic loop to include both psychophysiology and neurophysiology data and discussed how such loops could be applied to a range of applications from robotics to digital health and gaming. Fairclough discussed that the use of implicit adaptation is an under-utilised mechanism that should be explored further. This is in reference to how the use of physiological adaptation could eventually replace the need for individuals to self-report on their current state. Such a system would work on the basis of training a classifier on labelled data (i.e. self-reported, expected outcomes) to then adapt a system in real-time based entirely on a user's psychophysiological data.

Zander and Kothe (Zander & Kothe, 2011) explored the concept of providing system adaption without the conscious effort of the user, which is referred to as passive Brain Computer Interface (BCI). Zander and Kothe (Zander & Kothe, 2011) presented the theory that passive BCI could be used to improve the functionality of current BCI systems, by providing implicit data to the system that would usually be difficult to record or observe, such as EEG data. The creation of a passive BCI system is explored further by Zander et al. (Zander, Krol, Birbaumer, & Gramann, 2016), who created a passive BCI system that used a probing technique to complete task goals. Previous research ((S. Fairclough & Gilleade, 2012a; Pope et al., 1995)) has focused on creating adaptive systems that respond to spontaneous changes in neurophysiology or psychophysiology in response to stimuli. However, Zander et al. (Zander et al., 2016) developed a passive BCI system that was designed to elicit specific responses from the user, in order to use these responses to train the system. The system created in this work used a cursor that was presented to the user and a target position on screen. The user was informed that the experiment would end when the cursor reached the target on the screen, but they were unaware that their implicit responses to cursor movement would be used to determine the direction the cursor should take in order to reach the target position. The experiment would begin, and the cursor would move around the screen in a randomised order. Each time the cursor moved, the movement and the participants neurological response were recorded. These neurological responses were labelled depending on whether the cursor had moved towards or away from the target position. The movement of the cursor was designed to elicit responses from the user (hence the use of the phrase probing) so that eventually, the user's responses could be used in order to decide which movements the cursor should take in order to reach the target position.

Passive BCIs can be used either to achieve a targeted outcome (such as the study presented by Zander (Zander et al., 2016)) or to monitor a participants brain activity in order to make beneficial changes to a system. Passive BCIs implement cognitive monitoring to gather information about a user's cognitive state. Passive BCIs can be used either independently, or to enhance active BCI systems. The use of passive BCIs enables a novel input technique and enables the use of data which would not be possible without cognitive monitoring. Passive BCIs can be used to monitor users intentions, situational interpretations and emotional states (Zander & Kothe, 2011).

1.3.2 Neuroadaptive Gaming

1.3.2.1 Dynamic difficulty adjustment

Neuroadaptive gaming refers to games that can be implicitly or explicitly controlled by neurofeedback. One goal of neuroadaptive gaming is the real-time dynamic adjustment of game difficulty. This difficulty adjustment is called Dynamic Difficulty Adjustment (DDA) or Dynamic Game Difficulty Balancing (DGDB). The purpose of this adjustment is to optimise the level of challenge for the player in order to encourage the player to enter the flow state. Referring back to section 1.1 of this chapter, the flow state lies between the undesirable states of boredom and overload (see Figure 1), which should be avoided. The literature discussed in this chapter leads to the theory that the most effective distraction from pain will be found if we are able to find the optimal level of challenge for the player, which requires an understanding of exactly how the player is responding to the current level of challenge, and how their skills level corresponds to that level. Commercial games have DDA algorithms built into them that rely on game metrics in order to make adaptations. Commercial DDA generally modulates factors, such as game speed, pickup rate (the rate in which both shield objects, weapons or health kits) and player/Non-Playable Character (NPC) power. For example, if a player is playing a racing game and is in last place, DDA could increase the speed of the player (or decrease the speed of the NPCs) in order to allow the player to achieve more of the in-game goals (Hunicke, 2005). DDA allows for game adaptions that can be made in line with inter-individual differences. For example, game difficulty level can be reduced for players whose metrics indicate that they are less skilled than the average player. In this case, metrics such as current game score, could be used to determine skill level

(Jennings-Teats, Smith, & Wardrip-Fruin, 2010). Using DDA can allow for the creation of games that can adjust their difficulty based on players current performance, even if it differs to their historical performance. For example, if a player is feeling tired or bored, their game performance may be reduced compared to their expected performance, and therefore they may require difficulty adjustments. Adjusting game difficulty in line with intra-individual factors allows for a more sensitive adaptive system, which will not only provide the optimal level of challenge for each individual, but also provide the optimal level of challenge for one player in a variety of different mental states.

As previously mentioned, one important factor that must be considered for DDA is the time window required for data collection. Regardless of the type of data that is collected (i.e. game metrics, physiological data, neurophysiological data), it is important to ensure that the time window is both sensitive enough to determine current performance, but not so sensitive that adjustments will be too frequent and therefore disruptive. The speed of the game is an important factor to consider alongside the data-collection time-window. For example, a racing game may need to make more frequent adjustments than a slower-paced game, such as a puzzle game. It is important to remember that the purpose of employing DDA techniques it to provide an enhanced gaming experience, as opposed to a non-adaptive game, so careful considerations must be made during the development phase. The act of measurement and adjustment, based on in-game actions and metrics, may not be sensitive enough for the creation of an adaptive game, as the actions made in game are discrete and provide limited sampling points.

1.3.2.2 Neuroadaptive/Physiological Adaptation

The benefit of DDA is that games can be adapted in real time to provide a more immersive experience for a participant. However, DDA relies on the information provided by the game to be a reflection of both the players level of engagement and their overall opinion of the game. Using DDA means that there is no way to ensure that the correct difficulty modulation is chosen. However, there are alternative ways in which the difficulty of a game can be modulated, which may provide a more successful decision and outcome. Some studies (Chanel, Rebetez, Bétrancourt, & Pun, 2008, 2011; Fernández, Koji, & Kondo, 2017; Szegletes, Köles, & Forstner, 2014) have been carried out using physiological adjustment, wherein the difficulty of the game is modulated by the participants physical response to the gameplay experience. These changes can either be based on the conscious efforts of the participant, or their unconscious response to stimuli. For instance, Chanel et al. (Chanel et al., 2008) found that, through the deliberate induction of three mental states (boredom, engagement and anxiety), a classifier could be trained on peripheral physiological data with the intention of determining these mental states in real-time. The average classification rate achieved in this study was 53%. In 2011, this work was expanded (Chanel et al., 2011) to create an adaptive version of Tetris, which was controlled by Electroencephalogram (EEG) and peripheral signals, in order to maintain player engagement and prevent undesirable mental states during gameplay. The researchers found that mental states could be classified with an accuracy of 63% when both EEG and peripheral signals were combined. Additionally, in 2014, Szegietes et al. (Szegletes et al., 2014) measured EEG, eye movement, pupil diameter changes and electrocardiogram (ECG) in order to determine how close a player was to the flow state.

Although this study did not develop an adaptive game, it did outline a framework that could be followed for the future creation of adaptive games. In 2016, further work was carried out in by Ewing et al. (Ewing, Fairclough, & Gilleade, 2016), whereby an adaptive version of Tetris was created using EEG to maximise game engagement. This study used a variety of adaptive systems (conservative – moderate – liberal) to determine which adaptive system was more pleasing for the user. In 2017, Fernandez et al. (Fernández et al., 2017) created a neuroadaptive game using EEG and game performance data, which was classified in order to find the optimal level of game demand for the player (specifically to provide an enjoyable experience). The results of this study indicated that a successful adaptation of game difficulty could be made using EEG measures.

Neuroadaptive games require real-time processing of brain signals, and then use machine learning classifications to classify that information (for example, classifying engagement) in order to determine how changes can be made to make a game more immersive. The theory of using neurophysiology, as opposed to DDA, is that brain signals can be classified into low engagement and high engagement, which negates the issue of determining player engagement based purely on player performance. We theorise that neuroadaptive technology will enable us to create the most immersive gameplay experience, which can then be tested to explore the theory that a neuroadaptive game will provide a better distraction from pain than a standard game. The benefit of using technology such as neuroadaptation (as opposed to DDA/DGDB) is that player state, which may be more sensitive than game metrics alone, can be continuously quantified in real time, without requiring the user to perform tasks specifically designed to measure their mental state, such as completing a questionnaire.

Neuroadaptive games can be created for a variety of reasons. One example of a neuroadaptive game (Krol, Freytag, & Zander, 2017) takes in multimodal input from the user for the purposes of direct game control (replacing the need for mouse/keyboard input) and for player state monitoring. This game was based on an adaptation of the game Tetris and was controlled using eye-tracking techniques. The neuroadaptive element of this game related to the players cognitive state and was created as a passive system. Game speed was modulated by players level of relaxation – if classification indicated that players were relaxed (in opposition to being engaged) then the speed of the game and music was increased. The increase in game speed was implemented with the intention of encouraging a state of engagement. In addition to the modulation of game speed, Tetris pieces could be removed from the board after they fell if the system classified that the last movement was an error. i.e., if classification from the players brain indicated that they did not wish for the desired piece to fall into the Tetris grid as it did, then this piece was removed and replaced by a new random piece at the top of the field. The purpose of neuroadaptation in this case is both to add functionality to an already existing game (Tetris pieces are not usually removed from the field) and to create a more engaging gameplay experience. This research indicates that neuroadaptive elements can be created and applied to gaming technologies both to alter the gameplay in general (to improve or change a game) and for the purposes of encouraging desired mental states. The purpose of modulating attentional states could be purely to improve the gameplay experience and increase enjoyment – but neuroadaptive gaming systems can also be created to encourage enjoyment or engagement for the purpose of using a game as a distraction. The work presented in our studies uses neuroadaptive gaming to encourage immersion in order to distraction from a painful experience.

Our hypothesis that a neuroadaptive game could distract from a painful experience is built upon previous research relating to workload adaptive processes. Biocybernetic adaptation has been in development since 1994 when research (Pope et al., 1995) was carried out to determine whether the appropriate time to switch between pilot control and autopilot control could be determined by brain activity. Since 1995, research into workload adaptive processes has continued – with recent research including the use of neuroadaptive support to determine how reading speed and readability difficulty could be modulated to improve the reading and learning experience (Andreessen, Gerjets, Meurers, & Zander, 2020).

1.3.3 Creating a Biocybernetic Loop

In order to develop a closed biocybernetic loop-based system, interdisciplinary knowledge is essential. The creation of a biocybernetic loop requires a knowledge of both signal processing procedures, as well as an understanding of how neurophysiological data should be treated and analysed. Fairclough and Gilleade (S. Fairclough & Gilleade, 2012a) published a paper that developed a biocybernetic loop and discussed the imperative steps for proper development. According to this research, the first step in creating a biocybernetic loop is to ensure that the data that is collected is rooted in proved theory. Namely, that the sites from which data are collected can be proven to provide insight into the mechanisms of the brain that the adaptive system is aiming to manipulate. In order to ensure that the game is affective, targeting brain areas that have been thoroughly researched and have been proven to relate to the experiences of pain, distraction and attentional control as a whole is a necessity. Previous research that has already been discussed in this chapter has indicated that analgesic benefits of computer games are caused by the games' ability to perform as a distraction and activate the top-down attentional control system of the brain.

The second stage of creating a biocybernetic loop, as suggested by Fairclough and Gilleade (S. Fairclough & Gilleade, 2012b), is to ensure that the measures that are collected can deliver a sufficient representation of the user state. Again, previous research discussed in this chapter has indicated that neurophysiological measures can be used to create an adaptive system that provides more successful outcomes than non-adaptive systems. However, it is important that we ensure that the measures that are derived from raw neurophysiological data are providing a clear insight into brain activity. For this reason, we much consider a variety of neurophysiological measures that can be derived from the brain signal and determine the measures that can most accurately represent the state of the participant.

The third step in creating a biocybernetic loop is to ensure that the classification of participant data is accurate. For this reason, it is important to consider more than just the accuracy of each classification that is made by the system but also consider other factors, such as the error rate. It is important also to remember that there is a lag present in some types of neurophysiological data, which means that the timings of adaptation and then data collection and classification must be considered carefully to ensure that the data that is being classified is providing an accurate representation of the brains response to game adaptation.

The final stage of the creation of a biocybernetic loop is to ensure that the adaptation to the system that will occur has the desired effect on the user. It is important to ensure that, if changes are made to the game, these changes are relevant to the user and fulfil the goal of the system. For example, if classification determines the level of challenge to be too high, the adaptive system must make changes that are proven to reduce the level of challenge that is experienced. If the system cannot make these changes, then the biocybernetic loop will not be able to perform as intended.

Fairclough and Gilleade (S. Fairclough & Gilleade, 2012b) also presented a case study in their work that followed the creation of a biocybernetic loop. Their biocybernetic loop consisted of six stages, and the overarching purpose of the loop was to modulate game difficulty depending on players states. We find it pertinent to follow this model as we create our own biocybernetic loop, due to the similarities found between the two projects. The six steps as inspired by Fairclough and Gilleade are discussed in detail below:

1. **Conceptual Model –** The first step that should be followed is to develop a model of how the loop should work, which must consider both the psychological concepts that support the loop and the expected outcomes that the behaviour of the loop will have on the user. For this reason, we must ensure that we have i) a theory (supported by the literature) that we aim to influence or explore (in the case of this work, this theory is the top-down modulation of attention to prevent bottom-up influences) and ii) a consideration of how the mechanisms of the loop will influence the user (for this work, we expect that modulating game difficulty will allow us to identify the optimal level of difficulty for each player.)

2. **Psychophysiological Inference –** The second step that should be followed is to determine what data should be used to infer the psychophysiological theorems that we aim to explore. We aim to create a system that can modulate game difficulty based on the theory that a higher level of top-down attention towards a distraction task will reduce the perception of pain. For this reason, throughout the course of this work, we must explore previous research that will allow us to determine the areas of the brain wherein we can capture data that represents both the top-down and bottom-up attentional control systems.

3. **A Quantified Model of User State –** The third step that should be followed indicates that we must ensure that the data we collect can be used to quantify the users *actual* state. It is not enough for us to assume that a game is difficult, we must quantify this level of difficulty before we can ensure that the biocybernetic loop will work. For this reason, we must first create a non-adaptive game with varying levels in order to determine the levels that the user finds i) too easy, ii) too hard and iii) are more optimal. Quantification could be carried out via a series of questionnaires. However, since we are modulating game difficulty with the aim of distracting from pain, we must also ensure that we can induce, measure and reduce the experience of pain in a quantifiable way.

4. **A Real-Time Model of User State –** Stage four is to ensure that the user state can be classified in real time. Referring back to the work of Csikszentmihalyi (Csikszentmihalyi, 1975), there are three users states that must be considered when attempting to encourage the user to enter the flow state, including boredom, flow and overload/anxiety. In order to be able to model the user state in real time, we must have quantified data that represents these states. In order to classify user state in real time, we will use machine learning classification. However, this requires training data, which will be collected prior to the real-time adaptive study, in order to classify new real-time data. In our work, this means that the data that was collected during step 3 must be given a label to represent the user state at the time when data was collected, which can be used as training data. Then, new data that is received during the real-time system can be classified against this training data, to provide a real-time classification of current user state. This classification will then be used to inform the biocybernetic loop of required adaptations.

5. **Design of the Adaptive Interface –** Stage 5 concerns the exact nature of how the game adaptations should work. Fairclough and Gilleade (S. Fairclough & Gilleade, 2012b) note that, although overt adaptations may be immediately impactful, they can also draw attention to errors, such as an incorrect adaptation caused by an incorrect classification. For this reason, covert adaptations will be used both to ensure that potential errors are not obvious and off-putting to the user, and also to ensure that the users experience of immersion is not disrupted by drastic changes in the challenge level of the game. It is postulated that more frequent covert changes will provide a more immersive experience than less frequent overt changes.

6. **Evaluation –** Stage 6 is the final stage in the development of the biocybernetic loop and concerns making evaluations as to whether the adaptive game has been successful. The goal of our proposed adaptive game is to reduce the perception of pain by increasing engagement towards a distraction task. In order to determine whether the adaptive game has been successful in achieve this goal, we must also define a comparison condition. We understand from previous research that performing a task whilst experiencing pain can reduce the perception of pain, so it would not be sufficient to compare the adaptive game to a painful experience experienced without a distraction task. Once we have created both an adaptive game and a comparison condition, we must then also measure changes in pain tolerance and pain perception between these two conditions, in order to efficiently evaluate the efficacy of the biocybernetic loop.

Overall, following the frameworks discussed in this section will allow us to create a biocybernetic loop that can power an adaptive game. However, as discussed, it is important that we understand exactly how the creation of this loop should begin, i.e., which measures we will use to determine player state, and how we can monitor both the top-down and bottom-up attentional networks within the brain.

1.4 Aims and Objectives

The aim of this project is to create a neuroadaptive game that is designed to perform DDA in real-time, in order to distract a player from pain. Functional Near Infrared Spectroscopy (fNIRS) will be used to create this neuroadaptive game. In order to create a prototype such as this, there are a number of objectives that must be carried out:

1. Firstly, markers of top-down attentional regulation must be identified based on neurophysiological data. One of the main advantages of fNIRS data collection is that the data collection points can be chosen by the user. This means to say that an individual and bespoke montage can be set up for data collection to ensure that the relevant brain areas are being monitored. Previous research has allowed us to understand the patterns of neurophysiological activity we can expect to see when the top-down attentional control network is activated. As such, the first step in this study is to determine the areas of the brain where we can identify this activity.
2. Once we are satisfied that we have created an fNIRS montage that can accurately capture data relating to the top-down and bottom-up networks of attentional control, we next need to assess whether we can classify changes within the top-down and bottom-up networks. To do this, we must design situations with the intention of inducing different mental states (boredom/flow/overload; pain/no pain) and then use machine learning algorithms to assess the validity of classification of these data.
3. Thirdly, we must assess the relationship between pain tolerance and game difficulty. Previous studies have indicated that the use of a game as a distraction task can improve pain tolerance. However, we need to ensure that both the distraction task and the painful stimuli are sufficiently correlated with each other, before we can create a game that is designed to modulate these factors.
4. Once we are satisfied that we have found a significant relationship between pain and game difficulty, which can be classified using machine learning techniques, we must then collect a set of training data, which can be used to train a machine learning classifier to detect the attentional state of the player. This data must represent each of the three mental states that we expect to capture (boredom, flow and overload) to ensure that any of these data types can be classified in real-time.
5. It is important to consider that real-time neuroadaptive data is likely to be contaminated by noise. A noisy signal could have detrimental effects on the accuracy of the real-time classifications, and could lead to incorrect classifications, and therefore incorrect adaptations being made to the game. For this reason, we much research and define a signal processing pipeline that can work in real time, to treat the collected data prior to classification.
6. There are a variety of measures (both physiological and neurophysiological) that have previously been used to create adaptive systems, so it is important that these measures are thoroughly tested in our system. Although we expect that neurophysiological data will provide a more accurate classification of top-down attentional control that alternative measures, some alternative measures (such as game metrics and physiological measures like heart rate) are

easier and cheaper to collect, easier to process and more comfortable for the participant. For this reason, we find it pertinent to explore a variety of measures and determine the measures that can actually provide the most accurate classification rates.

7. Finally, after following these steps we expect to have a working prototype of a neuroadaptive game. However, this prototype will still need to be validated. It is important not only that this prototype can function as expected in a technical capacity, but also that the neuroadaptive system can be validated to determine whether it can provide a distraction from experimental pain. It is also important to explore whether a neuroadaptive game is more effective than a standard game at distraction the participants from pain.

1.5 Novelty

To our knowledge, there have been no previous studies carried out with the express purpose of creating a neuroadaptive game that will distract a participant from pain. Although video games have been used as distractions, and various different distractions have been compared to each other, we have not found any literature that uses real-time neuroadaptation in this way. Similarly, although there are a variety of studies that have explored the filtering of fNIRS signals, we have been unable to find any papers that outline an efficient, real-time fNIRS signal filtering protocol. Overall, our research has not indicated any papers that are in our field of real-time fNIRS processing for the purpose of real-time classification in order to distract participants from pain. We have also been unable to find any papers that deal with the issues of communicating fNIRS signals in real time between two machines, which is an integral part of our data processing pipeline, and an overall aim of our research. Overall, although the backbone of our research is based around neuroadaptation and the experience of immersion, which have been explored in the literature, the overall purpose of the project and the underlying mechanisms that will be created to facilitate this goal are novel. The novel contributions of the project will be as follows:

1. Using fNIRS to study attention (specifically in relation to pain) – to our knowledge, there is no previous research which uses fNIRS technology to examine the effects of pain specifically in relation to attentional control. We aim to determine which sites and features can be gathered from fNIRS data which directly relate to pain and attention, both as individual experiences and their effects on each other.
2. Creating a real-time signal processing pipeline – research indicates that a real-time signal processing pipelines has not been created specifically to process fNIRS data relating to attentional control and the experience of pain. Although software exists which can filter and process fNIRS data, it is our intention to create a bespoke system which can collect, process, visualise and make adaptations based on real-time data.
3. Creating an adaptive game that combines fNIRS and game metrics in real-time, via machine learning – Neurophysiological data and game metrics have been used previously to make adaptive gaming technology. However, it is our intention to find a combination of both data types which can create and effective adaptive gaming technology. We believe that a

combination of neurophysiological and game data will enable us to ensure that our adaptive pipeline is as sensitive and effective as possible – providing a more adaptive experience for the end user.

The research will be valuable to the fields of both fNIRS research and neuroadaptive technology. We believe our data processing pipeline will be transferable – meaning that the same filtration and feature creation techniques could be suitable for other fNIRS based BCI's – even those that do not relate to gaming or the pain experience. Similarly – we believe that the research we will undertake to identify metrics which relate to attentional control and the pain experience may be useful in alternate projects, which focus on these areas. Overall, we believe that our research could be used to further expand the fields of neuroadaptive technology, with a specific focus on the use of real-time fNIRS.

1.6 Summary

This chapter has presented an introduction to the research project, aims and objectives. A literature review relating to attentional control, the experience of pain, BCI and neuroadaptive technology has been undertaken. Chapter 1 serves to introduce the reader to the background and purpose of this research project, which will be the underlying focus of the remainder of this work.

Chapter 2 - General Methods

2.1 Introduction

This chapter describes the methods and procedures that were common to the studies performed in this thesis and are described in this section with scientific background. Three studies were carried out during the course of this research project. Although there were some modifications to each independent study, the three studies all followed a similar protocol. The purpose of the first study was to determine how game difficulty level would affect a participant's experience of pain. One video game, with four levels of difficulty, was used in this study. Each participant played each level of difficulty twice – once whilst experiencing pain, and once without the experience of pain.

The purpose of the second study was twofold, firstly to determine whether we could find statistical significance in the response to game difficulty over three levels, and secondly, to determine how connectivity measures (which will be discussed in further detail in Chapter 4) could influence our results.

The purpose of the third study was to implement a neuroadaptive game and determine whether the neuroadaptive game provided more of a distraction from pain to a participant than a non-neuroadaptive game.

The structure of this chapter will explore the measures (fNIRS, ECG, subjective questionnaires), protocols (Cold Pressor Test (CPT), games) and the machine learning techniques that were used for data classification. Figure 3 below indicates the protocol that was followed throughout the course of these studies. The purpose of each of the three studies was to determine if, and how, changes in game difficulty or function would affect a participant's experience of pain. For this reason, it was important to ensure that we measured the participants pain tolerance without any distraction at all, as well as with each different distraction. It is also important when collecting fNIRS data that a baseline is collected, to allow for the application of filters (Baker et al., 2014). In order to thoroughly test our theories, the following protocol was enacted to ensure that data was collected from the participant during each condition – ensuring that multiple distraction tasks could be compared both to each other, and to a participant's standard pain tolerance level.

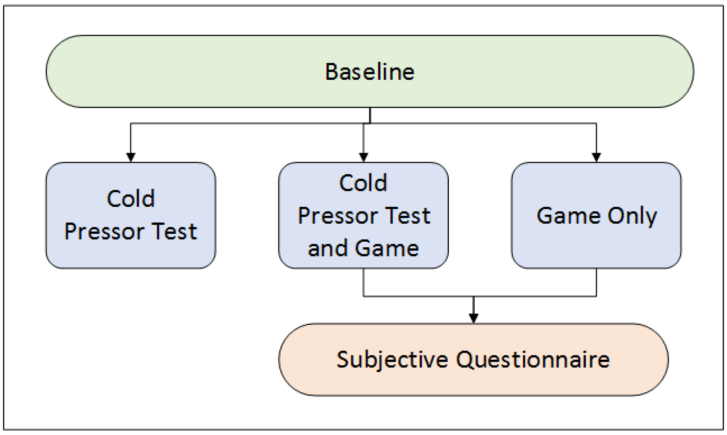


Figure 3 - Protocol followed for studies 1-3

2.2 Data Collection Techniques

2.2.1 Functional Near Infrared Spectroscopy

Functional Near Infrared Spectroscopy (fNIRS) is a functional, non-invasive measure of brain activity. fNIRS technology development began in 1977 and its first use in a clinical experiment was in 1985, to study the cerebral oxygenation of new-born infants (Ferrari & Quaresima, 2012). fNIRS is a light-based technology, which relies on the transparency of human tissue at certain frequencies of light. fNIRS systems work via a series of sources and detectors that are positioned to match a desired montage across the cortex, to target specific regions. Once a montage has been designed, a series of sources and detectors are placed on a head cap to enable data collection. An example of the sources (blue) and detectors (yellow) can be seen below in Figure 4.

Figure 4 - An example of the sources and detectors used in an fNIRS montage

fNIRS sources emit near infrared light (NIR), which the skin, tissue and bone of the head are mostly transparent to. Oxygenated and Deoxygenated haemoglobin are strong absorbers of NIR light, specifically at frequencies of 700-900 nm. This absorption means that activity in relevant brain areas can be measured as the NIR light is emitted through the source and received back via the detector. Light is transmitted and received between sources and detectors in a banana-shaped curve, which can be seen below in Figure 5, taken from Naseer and Hong 2015.

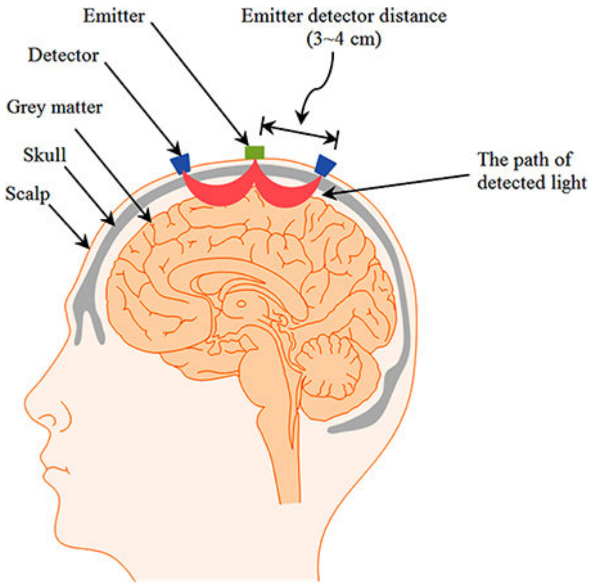


Figure 5 - Figure displaying the banana shaped light curve that is transmitted and received by the fNIRS hardware

The difference in light that is emitted versus the light that is received back is dependent on the amount of haemoglobin that is present in the measured area of the brain. These fluctuations in light depend on the current brain activity. For example, if a particular area of the brain is activated, then there will be a higher volume of blood in this area when compared to the same area of the brain in a non-activated state. Measuring fluctuations in brain activation allows for the determination of the areas of the brain that are activated, which in turn can be used to determine specific processes that may be occurring within the brain.

The optimal spacing for sources and detectors is 3 cm, which allows the NIR light to sufficiently absorb before it is scattered and received. If haemoglobin is oxygenated (HbO), it absorbs more light, and therefore less light is scattered and received by the detector. Measuring the absorption of NIR light allows for the determination of oxygenation in different areas of the brain. Activation in areas of the brain is referred to as neurovascular coupling, an example of which can be seen below in Figure 6[2], which is categorised as a simultaneous increase in oxygenated haemoglobin (HbO) and decrease in deoxygenated haemoglobin (HbR) (Berger, Horst, Müller, Steinberg, & Doppelmayr, 2019; Ferrari & Quaresima, 2012).

[2] https://www.gowerlabs.co.uk/fnirs

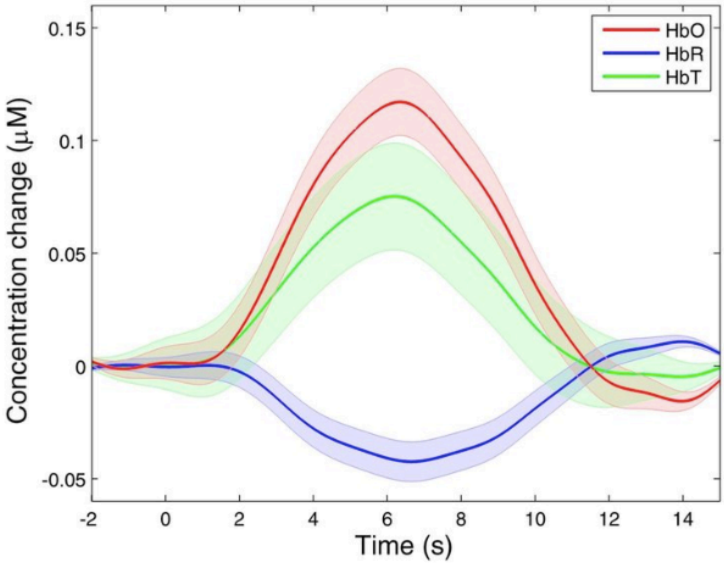


Figure 6 - Hemodynamic response to neurovascular coupling[3]

fNIRS technology is not without its weaknesses, which mainly relate to noise within the signal. As expected, when the head moves, the blood in the brain moves also. Since fNIRS technology relies on blood oxygen level dependant (BOLD) responses, these movements can cause noise. Similar to EEG, movement or slippage of the fNIRS optodes can also create a noisy signal. fNIRS is also susceptible to noise that stems from heart rate changes and respiration fluctuations. Although some of this noise can be removed with relative ease using filters (Bauernfeind, Wriessnegger, Daly, & Müller-Putz, 2014), some noise can be more difficult to remove. One of the most prominent issues that can occur if noise is not sufficiently removed from an fNIRS signal is the presence of false positives and false negatives in the data. The terms false positive and false negative are used to describe what can happen when noise disrupts an fNIRS signal in such a way that it appears that neurovascular coupling is (false positive) or isn't (false negative) happening (Tachtsidis & Scholkmann, 2016).

Although fNIRS is designed to record data relating to neuronal activity, other variables can be collected within this data, which can obfuscate the signal. Variables that can affect the clarity of the fNIRS signal are referred to as noise. Typically, fNIRS signals contain data relating to systemic, neuronal, cerebral and non-cerebral variables. Previously, it has been assumed that fNIRS signals relate only to task-evoked cerebral and neuronal activity, i.e. activity that relates only to a genuine neurovascular response to a specific task or stimuli. However, fNIRS data actually contains information relating to non-task-evoked systemic and extracerebral data, as well as task-evoked cerebral and neuronal activity. Non-task-evoked data refers to fNIRS data that does not relate specifically to the task at hand, for example, mind wandering or distractions that the participant experience. Extracerebral data refers to data that is collected by the fNIRS sensor from the head/scalp rather than the brain, for example, brow furrowing or muscle movement. Systemic data refers to changes in blood activity that relate to changes in heart rate or respiration which, although may be task evoked, do not accurately indicate neurovascular coupling. An example of how these variables can affect the fNIRS signal can be

[3] https://www.gowerlabs.co.uk/fnirs

seen below in Figure 7, which is taken from (Tachtsidis & Scholkmann, 2016). It is important to remove as much noise as possible from the fNIRS signal to ensure that false positives and negatives are not misrepresented as neurovascular coupling.

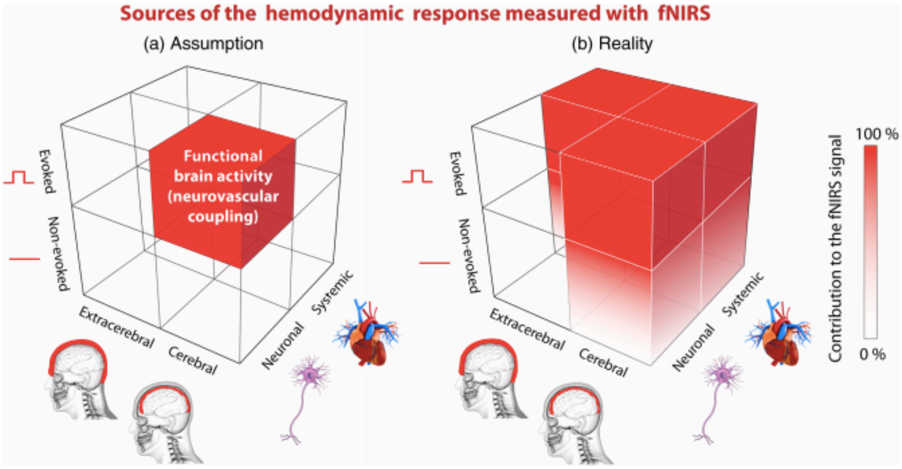


Figure 7 - Variables which can cause false positive and false negative indicators of neurovascular coupling in an fNIRS signal (Tachtsidis and Schlokmann 2016).

fNIRS has also previously been used in real-time to control Brain Computer Interface (BCI) systems. In 2007, three papers reported that fNIRS had been used to control a variety of applications as a replacement for hardware control. For instance, Coyle et al. (Coyle, Ward, Markham, Ward, & Markham, 2007) used fNIRS data gathered whilst a participant imagined squeezing a ball to control a yes/no system. Whilst Siteram et al. (Sitaram et al., 2007) used fNIRS data to distinguish between physical and imagined movement and achieved a classification accuracy of 80%. fNIRS data has also been used in medical trials concerning Amyotrophic Lateral Sclerosis (ALS) with a success rate of 70% for non-locked in patients and 40% for locked in patients. The ALS study concerned answering ‘yes’ via performing a cognitive task and answering ‘no’ by performing no task (Naito et al., 2007). The tasks reported in this study were related to explicit control over a system using brain activity. However, other studies have explored the use of fNIRS data to implicitly control systems.

2.2.2 Electrocardiogram

In order to explore whether game demand could be classified by heart rate (HR) data, HR data was collected using a Zephyr BioHarness device, which collected data at a sampling rate of 250 Hz. This device was fitted to an elasticated strap and was worn by the participant under their clothing at the centre of their chest. This device recorded EEG data which was later converted into a representation of beats per minute (BPM) for data classification. EEG data was collected for Study 1 only.

2.2.3 Accelerometer

fNIRS data is known to be sensitive to noise relating to head movement, which can obscure the fNIRS signal and prevent a true representation of brain activity from being collected. One way in which this head movement data can be filtered out of the fNIRS signal is by monitoring head movement and building a personalised filter for each data file based on a combination of head movement data and artefacts contained within the fNIRS data. For this reason, an accelerometer was used to collect head movement data throughout Study 1.

2.3 Data Collection and Pre-Processing

A data pre-processing pipeline has been created, using a variety of filters and algorithms, which were applied to the raw data (see Figure 8). These filters were applied to ensure that the signals were free from artefacts, which could affect the results.

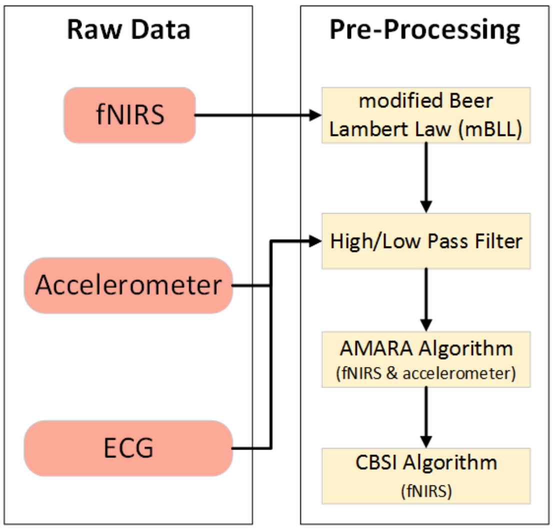


Figure 8 - Data pre-processing pipeline

During the three experiments carried out throughout the course of this research, fNIRS data was collected using the Artinis Oxymon Mk III system. This system consists of two control boxes, into which both the transmitter and receiver cables are attached. The receiver and transmitter cables consist of a connection to attach them to the control box on one end, and optodes on the alternate end. The optodes are attached to the optode holders, which are placed in a montage onto the fNIRS cap, which is worn on the participants head for the duration of data collection. The collection of data is managed by the Artinis software Oxysoft, which enables data to be collected, stored, and then exported for analysis. Examples of the Artinis Oxymon system, and the fNIRS cap, can be seen below in Figures 9-10[4].

[4] https://www.artinis.com/oxymon).

Figure 9 - Artinis Oxymon fNIRS system

Figure 10 - fNIRS Cap

Before fNIRS data can be analysed, pre-processing must be carried out in order to provide a filtered signal that is representative of Oxygenated and Deoxygenated haemoglobin. These processes are discussed in detail below.

2.3.1 Modified Beer Lambert Law (mBLL)

Data is received from the Oxymon system in the form of optical density (OD) data, which determines the change in light absorption between a transmitter and receiver. In order to establish how the OD data relates to activation within the brain, it must first be converted to a representation of Oxygenated (HbO) and Deoxygenated (HHb) haemoglobin. This conversion is performed through the application of the modified Beer Lambert Law (mBLL). One important element of the mBLL is the determination of differential pathlength factor (DPF), which is dependent on wavelength, age of participant and tissue type. The mBLL is displayed below in equation 1:

$$\Delta \mathrm{c} = \frac{\Delta \mathrm{OD}_{\lambda}}{\varepsilon_{\lambda} \cdot L \cdot DPF} \tag{1}$$

In equation 1, $\Delta \mathrm{c}$ is equal to the converted fNIRS signal, $\Delta \mathrm{OD}_{\lambda}$ refers to the oxygen-independent optical loss that occurs due to scattering and absorption. ε_{λ} refers to the extinction coefficient, L refers to the distance (in cm) between the light entry and exit points, and DPF refers to the differential path length factor.

Algorithm 1 presents the mBLL, which is used to convert raw optical density (OD) data to oxygenated haemoglobin (HbO) and deoxygenated haemoglobin (HHb) data:

Algorithm 1. Application of the modified Beer Lambert Law

```
Data: opticalDensity
Result: Delta_HHb, Helta_HbO

cy = 2;
for mBLLLengthMarker = 1:length(opticalDensity)-1
    Delta_OD_858(mBLLLengthMarker,1) =  (opticalDensity(cy,1) -
opticalDensity(1,1)) / (Distance*DPF);
    Delta_OD_764(mBLLLengthMarker,1) =  (opticalDensity(cy,2) -
opticalDensity(1,2)) / (Distance*DPF);
    cy = cy+1;
end
for mBLLLengthMarker = 1:length(opticalDensity)-1
    Delta_HbO(mBLLLengthMarker,1) = (Inverses(1,1) *
Delta_OD_764(mBLLLengthMarker,1)) + (Delta_OD_858(mBLLLengthMarker,1) *
Inverses(1,2)) ;
    Delta_HHb(mBLLLengthMarker,1) = (Delta_OD_764(mBLLLengthMarker,1) *
Inverses(2,1)) + (Delta_OD_858(mBLLLengthMarker,1) * Inverses(2,2));
End
Delta_HHb = Delta_HHb * 1000;
Delta_HbO = Delta_02Hb * 1000;
```

Algorithm 1 presents the process for converting raw OD data to HbO and HHb data, following Equation (1) in Chapter 2. The first step in performing the calculation is to define fixed variables, which will remain consistent for each individual's dataset, including:

(a) ***HbO and HHb Extinction Coefficients*** – These refer to the speed at which a substance absorbs light, depending on the wavelength of said light (Zhao, Qiu, Sun, Huang, & Li, 2017). The fNIRS measurement Wavelengths are 858nm and 765nm, meaning that the extinction coefficients are 1.6121, 0.6345 and 0.7949, 1.198, respectively, for each wavelength (Cope, 1991).

(b) ***Extinction Coefficients Inverse*** –The inverse of (a).

(c) ***Source-detector separation distance*** – This defines the measurement between each source and its corresponding detector. In this study, this value was fixed at 3.1cm.

(d) ***Differential Pathlength Factor (DPF)*** – This variable is different for each participant and relates to the change in absorption. This can be affected by age and source-detector separation (the distance between the source optode and its detector pair). Prior to data collection, DPF is calculated by the Artinis Oxymon software, per each participant (Kamran, Mannann, & Jeong, 2018; Scholkmann & Wolf, 2013).

Following the definition of these fixed variables, Equation (1) (Chapter 2) is solved to provide mBLL transformed data. Line by line, this script is executed as follows:

- Line 1 instantiates a marker number, which is used to ensure that the correct datapoint is selected from the file.
- Line 2 instantiates a for loop, which runs for the duration of the *opticalDensity* file, minus one. This loop allows the algorithm run continuously over the entire file.
- Line 3-4 defines the value of *Delta_OD_858*, or the first datapoint which relates to light at the frequency of 858nm (which resides in column 1). This datapoint is called via the value of *cy*, which starts at 2. The value *Delta_OD_858* has the sum of the first value in the file *(opticalDensity (1,1)*) divided by the *Distance * DPF* subtracted from it.
- Similarly, Line 5-6 defines the value of *Delta_OD_764*, the first datapoint, which relates to light at the frequency of 764nm (which resides in column 2). Again, the value *Delta_OD_764* has the sum of the first value in the file (*opticalDensity (1,1)*) divided by the *Distance * DPF* subtracted from it.
- Line 7 advances the value of *cy* by one, which ensures that the loop works through the *opticalDensity* file continuously.
- Line 8 ends the loop now that the first stage of the transform has been completed.
- Line 9 instantiates a second loop, which selects the value of *mBLLengthMarker* and advances this value by 1 each time the loop runs. This ensures that the correct value is selected each time the loop runs, as the algorithm is applied to each data point individually.
- Lines 10-12 define the value of *Delta_HbO*. This value is determined via multiplying the value of Inverses (1,1) by the sum of the current datapoint of *Delta_OD_764* added to the sum of the current datapoint of *Delta_OD_858* multiplied by the value of Inverses at (1,2).
- Lines 13-14 define the value of *Delta_HHb*. This value is determined via multiplying the value of the current datapoint of *Delta_OD_764* multiplied by the value of *Inverses (2,1)*, added to the sum of *Delta_OD_858* at the current datapoint multiplied by the value of Inverses at (2,2).
- Line 15 ends the second loop in the algorithm, once the algorithm has been applied to all datapoints.
- Lines 16-17 multiply the currents values of Delta_HHb and Delta_HbO by 1000 to find the final values.

Figure 11 below displays both raw OD data, whilst Figure 12 displays HbO and HHb data following the mBLL transformation.

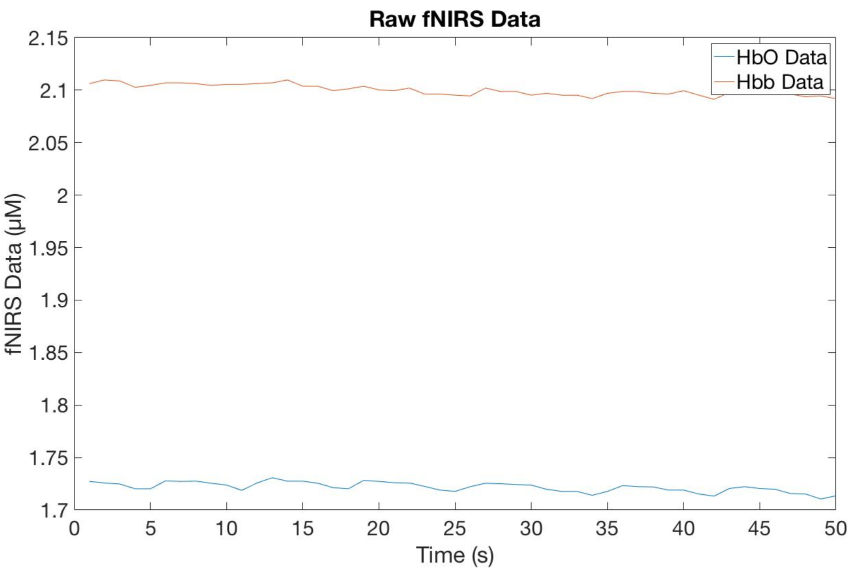


Figure 11 - Representation of raw OD data

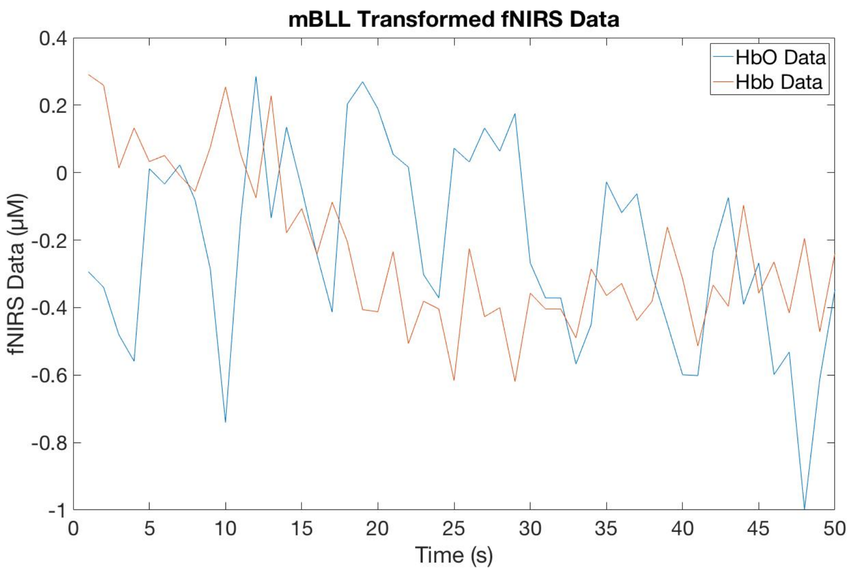


Figure 12 - Representation of HbO and HHb data after the application of the mBLL

2.3.2 High and Low Pass Filtering

Once the data has been converted to represent HbO and HHb, filtering is required in order to remove noise from the data. fNIRS data pre-processing is often conducted using specifically designed analysis packages, such as NIRSPM[5] and HOMER[6], or the software package that is used for data

[5] https://www.nitrc.org/projects/nirs_spm/

[6] https://homer-fnirs.org/

collection. For example, data processing of the Oxymon fNIRS data is usually carried out using the Oxysoft package provided, which are both supplied by Artinis. Although these filtering methods are validated and appropriate, the use of supplied filtering techniques has created a gap in the literature wherein the reporting of specific filtering techniques is often unreported. There are typically two types of noise that are present in fNIRS data, which require removal. The first of these being hemodynamic response, which is unrelated to neurovascular coupling. This noise is typically caused by fluctuations in heart rate, blood pressure and respiration. It is expected in coming years that the standard for removal of such noise will be applied via the use of short separation channels, which measure only hemodynamic response. This data can therefore then be used to remove such noise from the standard long channel fNIRS data, which is recording data related to genuine task-evoked response (Pfeifer, Scholkmann, & Labruyère, 2018).

Previous research that has been conducted using the Artinis Oxymon fNIRS system has been filtered using a combination of low and high pass filtering (Koenraadt, Roelofsen, Duysens, & Keijsers, 2014). High pass filters remove noise above a specified frequency, whereas lowpass filter remove noise below a specified frequency. This ensures that only data lying between these two frequencies remains. Typically, in fNIRS research, filters that retain data between 0.1 and 0.4 Hz are considered to remove a large amount of physiological data from the fNIRS signal, without distorting the signal (Naseer & Hong, 2015). There are a variety of filter types that can be applied to fNIRS data (e.g. Butterworth, Chebyshev, elliptical) but there have been no reports on any particular filter type being superior to another (Naseer & Hong, 2015). In this research, we have applied a 4th order Butterworth filter (Selesnick & Burrus, 1998) to our fNIRS data using MATLAB R2018B 9.5.0.944444. The cut-off frequencies were 0.3 and 0.1 Hz respectively. The High and Low pass filtering algorithm is displayed below in Algorithm 2:

Algorithm 2. Application of high and low pass filters

Data: baselined_fNIRS

Result: fNIRS_Filtered

```
fcl = 0.3;
fch = 0.1;
fs = 10;
% Lowpass filter
[filter_B,filter_A] = butter(4,fcl/(fs/2),'low');
fnirs_Filtered = filter(filter_B,filter_A,baselined_fNIRS);
% Highpass filter
[filter_B,filter_A] = butter(4,fch/(fs/2),'high');
fnirs_Filtered = filter(filter_B,filter_A,fnirs_Filtered);
```

In Algorithm 2, `fcl` refers to the low cut-off frequency and `fch` refers to the high cut-off frequency. This means that any data outside of the limits of these two frequencies should be considered noise and removed from the signal. `Fs` refers to the sampling rate of the fNIRS data, which in this case was 10Hz. `Butter` refers to the Butterworth filter that was applied to the data. An example of the effects of Algorithm 2 can be seen below in Figure 13.

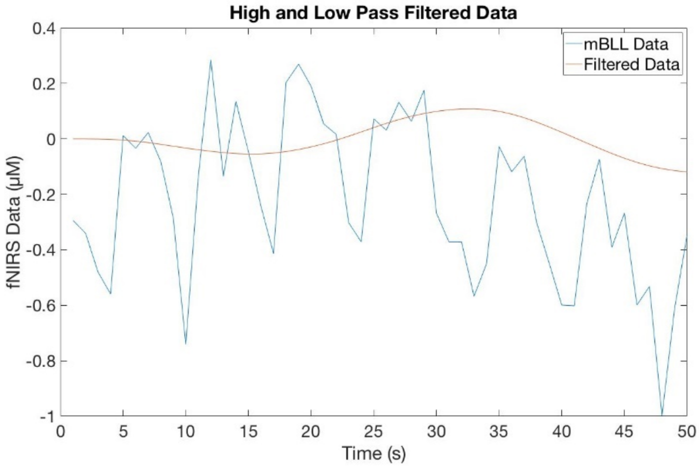


Figure 13 - Representation of raw and Bandpass filtered fNIRS data

2.3.3 Acceleration Based Movement Artefact Reduction

Artefacts relating to head movement are commonly found in fNIRS signals, which causes a change in blood flow to the brain. These changes can appear to represent changes in activation if they are not removed from the signal. The Acceleration Based Movement Artefact Reduction (AMARA) filter (Metz, Wolf, Achermann, & Scholkmann, 2015) detects periods of movement within an accelerometer signal and then compares these periods of movement to the fNIRS data. Where it is found that the moving standard deviation (MSD) of the fNIRS signal has changed considerably during the same period of time that movement has been detected, within the accelerometer signal, these segments of fNIRS data are marked as artefact segments. Segments where the MSD of both the accelerometer and fNIRS data have no significant deviation are marked as acceptable segments. Reconstruction of artefact segments uses forward and backward baseline adjustments and interpolation to reconstruct the entire signal, with the movement artefacts corrected. Acceleration data was collected for Study 1 only. Although AMARA provided to be effective in the post-hoc data filtration during Study 1, the algorithm performs best when it has a considerable time window of data. AMARA accounts for errors in the data caused by movement artefacts by interpolation of non-noise contaminated segments of the signal – although this is appropriate for post-hoc data analysis, we felt that applying this technique to real-time data (which has a considerably shorter timeframe) may have misrepresented the data in a real-time situation. In order to maintain the integrity of the data, this algorithm was not considered for Studies 2 and 3.

2.3.4 Correlation Based Signal Improvement

Head movement artefacts are expected in fNIRS data, especially during a gameplay scenario wherein the participant may move their head involuntarily as they become involved in the game. For this reason, the Correlation Based Signal Improvement (CBSI) function was applied to the fNIRS data (Cui, Bray, & Reiss, 2010a). This function finds points in the data where a positive correlation is found, which represents noise, as opposed to a negative correlation, which is expected from neuronal activity, and corrects the signals at this point to provide the expected negative correlation. The CBSI algorithm was validated via a study that monitored the HbO and HHb responses in the brain when head movement was deliberately induced. The study found that larger head movements caused a stronger positive correlation in fNIRS data than smaller head movements, and that this correlation was not equal across all channels. The head movement study also indicated that head movement noise occurs outside of the range of more commonly filtered fNIRS noise (such as autonomic fluctuations), which means that frequency-based filters may not be suitable for removing head movement noise. A further validation study found that the CBSI algorithm was suitable for removing both small and large amounts of noise, without degrading the overall quality of the data. The equations for the CBSI algorithm are as follows in equations 2, 3 and 4:

$$\alpha = \sqrt{\frac{\Sigma \mathrm{x}^2}{\Sigma y^2}} = \frac{std(x)}{std(y)} \tag{2}$$

The first step in performing the CBSI algorithm is to calculate α, which refers to the ratio of the standard deviation of HbO and HHb. In equation 2, $\Sigma \mathrm{x}^2$ refers to the sum of x squared, where x represents the Oxygenated haemoglobin data. Σy^2 refers to the sum of y squared, where y represents the Deoxygenated haemoglobin data. When α is calculated, the next step is to calculate the transformed HbO and HHb data. HbO is calculated via equation (3), and HHb is calculated via equation (4):

$$xo = \frac{1}{2}(x - \alpha y) \tag{3}$$

$$yo = -\frac{1}{\alpha} xo \tag{4}$$

In Equations 3 and 4, xo and yo refers to HbO and HHb respectively, and α refers to the value of α that was calculated in Equation 2. The CBSI algorithm was also applied using MATLAB, the process for which can be seen below in Algorithm 3:

Algorithm 3. Application of the CBSI algorithm

```
Data: oxy, deoxy
```

Result: oxy_CBSI, deoxy_CBSI

```
alpha = std(oxy)./std(deoxy);
oxy_CBSI = (oxy - (alpha * deoxy)) /2;
deoxy_CBSI = -(1/alpha * oxy(:,a))
```

An example of a movement artefact present in fNIRS data can be seen below in Figure 14. CBSI corrected data can be seen below in Figure 15.

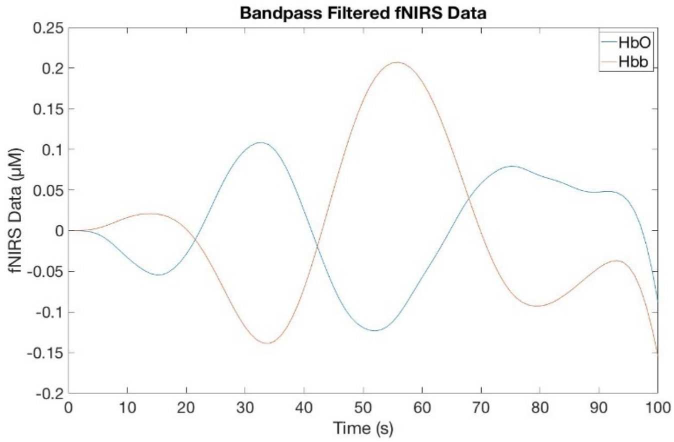


Figure 14 - Representation of bandpass filtered fNIRS data

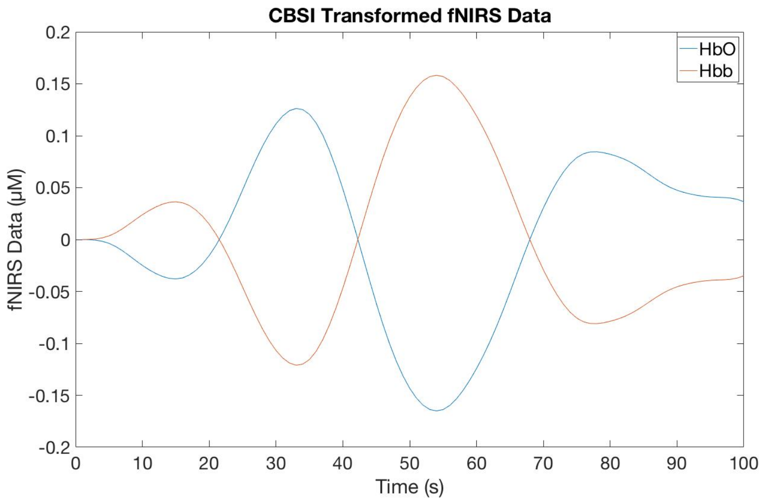


Figure 15 - Representation of CBSI transformed fNIRS data

One benefit of the two techniques that have been chosen for fNIRS data pre-processing (Bandpass/CBSI) is that they do not require a complete signal in order to successfully filter the data. This means to say that filtering can be performed either on epochs of data or the entire data stream, whilst maintaining the fidelity of the data itself. Filters such as this can therefore be used either post-hoc or in real-time, where, during a real-time filtering protocol, the filters can be applied to data epochs of any length for real-time filtering. A particular benefit of the use of the CBSI algorithm is that tHb can be calculated prior ot the application of the algorithm, which preserves the pattern of the date. CBSI makes assumptions about the expected shape of the data which may not necessarily be true in all cases – if only Oxygenated and Deoxygenated data were considered, there is potential for data loss which would reduce classification accuracy. However, using the CBSI algorithm alongside the non-CBSI treated tHb data enables us to ensure that any informative data that may have been corrected in the HbO and HHb is still included in the tHb data.

2.3.6 Cold Pressor Test

The Cold Pressor Test (CPT) is frequently used to induce experimental pain (Bullinger et al., 1984; Lovallo, 1975; Mitchell, MacDonald, & Brodie, 2004; Nouwen, Cloutier, Kappas, Warbrick, & Sheffield, 2006). During the CPT, the participant is asked to immerse a limb into the Cold Pressor (CP) stimulus tank. Typically, either the foot, hand or arm can be immersed. For the purpose of our studies, participants needed to use both hands to play the games, so were asked to immerse their foot into the stimulus tank. A typical CPT would use a water temperature between 2 and 8 degrees Celsius. The colder the water temperature, the quicker the participant will begin to experience pain. During Study 1, the game that the participant played lasted for three minutes. As such, it was determined that the CPT temperature would be set to 2 degrees Celsius. Setting the water temperature for 2 degrees ensures that the participant will experience pain regardless of the short gameplay session. The maximum amount of time that the participant was allowed to immerse their foot during the CPT was three minutes. Participants were informed that they could remove their foot from the stimulus tank whenever the pain became unbearable, regardless of the length of immersion.

Two Cold Pressor systems were used throughout the course of the three studies. The first system consisted of a large control tank, which had both an inlet and outlet valve, that was filled with ice and water. Water was passed between this control tank and a smaller stimulus tank via these pipes. The purpose of using both a control tank and a stimulus tank was to ensure that the water in the stimulus tank did not deviate from 2 degrees Celsius. Although this system was suitable for experimentation purposes, it was not reliable for continuous use due to issues with the internal pump system. For this reason, Studies 2 and 3 were carried out using one container, which consisted of two partitions. One partition was filled with water, and the participant would place their foot into this side of the tank. The other partition was filled with ice and water, and a pump was fitted to the inner wall to ensure a continuous flow of water, which ensured that the temperature of the participant's foot did not increase the overall temperature of the water. A thermometer was also fitted to this apparatus to ensure a consistent temperature. If required, additional ice could be added to this tank to ensure the water did not become too warm. Figure 16 below illustrates an example of the CPT that was used.

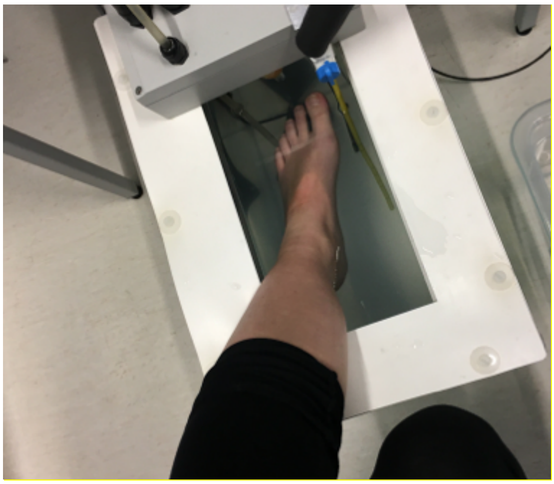


Figure 16 - CPT stimuli tank in use by a participant

This pain protocol was designed to measure pain tolerance, so participants were instructed to remove their foot from the water as soon as it become unbearable for them. This time was recorded by the experimenter as a measure of pain tolerance.

There are two main benefits which led us to choose to use the CPT for our experimentation. The first benefit is the quick recovery time and lack of extended implications of cold pressor pain. Once the participant has removed their foot from the water and the limb has had time to return to body temperature, the cold pressor test can be carried out again without concern for numbing of the limb or other factors which could affect the efficacy of the test. Thermal pain stimulation would also have the same benefits as the cold pressor test in terms of recovery time and repeated exposure. The cold pressor was chosen for our work as an alternative to thermal pain because the cold pressor provides a subjective measure of pain tolerance. Although pain perception was measured using the Visual Analogue Scale, this is a retroactive response and may not be reliable. Pain tolerance is an immediate response and may provide a more through indication of pain response, An immediate response may provide a more accurate indication as to how changes to the distraction task have affected pain tolerance and perception.

2.4 Subjective Questionnaires

A variety of questionnaires were used throughout the experiments to rate participant's pain experience and perceived experience of immersion, task load and motivation. These questionnaires were completed following each game condition, and each pain condition.

2.4.1 Immersive Experience Questionnaire

The Immersive Experience Questionnaire (IEQ) relates to participants overall feelings of immersion (Jennett et al., 2008). It consists of 32 questions, which are scored on a 5-point Likert scale from 1 (Not At All) to 5 (A Lot). Questions related to participants enjoyment of the game, as well as considering how involved the participant felt, and whether they felt as though they were inside the game as opposed to an external player. The IEQ ends with a question that asks participants directly to rate their feelings of immersion during gameplay. The IEQ can be seen in its entirety in Appendix 1. An example of some of the questions asked in the IEQ are as follows:

- *To what extent did you feel that the game was something you were experiencing, rather than something you were just doing?*
- *To what extent did you feel motivated while playing?*
- *To what extent did you feel emotionally attached to the game?*
- *To what extent were you interested in seeing how the game's events would progress?*

2.4.2 Visual Analogue Scale

The Visual Analogue Scale (VAS) was used as a measure of pain perception and consists of a 10-point scale, where participants can mark their level of pain on a line at any point between 0 (No Pain) to 10 (Unbearable Pain) (Gould, Goldstone, Kelly, & Gammon, 2004). The VAS can be seen in Appendix 2.

2.4.3 NASA Task Load Index

The NASA Task Load Index (TLX) is a 10-point Likert scale, with scores ranging from 1 (low) to 10 (high). It consists of 6 questions relating to mental demand, physical demand, temporal demand, performance, effort and frustration that the participant felt when playing the game (Hart & Staveland, 1988). The TLX can be seen in its entirety in Appendix 3.

2.4.4 Motivation

The Motivation scale is a 5-point Likert scale, with scores ranging from 0 (Not At All) – 4 (Extremely), which was used as a measure of intrinsic motivation. This scale consists of 8 questions relating to the participants thoughts and feelings both during and after the gameplay experience (Matthews et al., 1999). The Motivation scale can be seen in Appendix 4.

2.5 Games

Over the course of this project, Onteca developed two racing games for use in our studies. The first game that was used (during Study 1) was Space Ribbon. Space Ribbon exists as a commercial game, but Onteca provided us with a software development kit (SDK) that we used to vary the parameters of the game specifically for our experiments. It is important for the project that we used games that represented the commercial sector, so that future iterations of the project could use a variety of commercial games if this would better suit the interest and skill of the player. Space Ribbon and Ribbon Rush were specifically designed so that consistent levels of difficulty could be created and maintained throughout the entire testing phase. This ensured that each participant would experience the exact same level of difficulty as other participants, regardless of their skill level. The games were played on a MacBook Air using a Sony PlayStation 3 controller.

2.5.1 Space Ribbon

Space Ribbon is a zero-gravity, space themed racing game. Players compete against Non-Playable Character (NPC) cars (between 2 and 10 cars per race) in an effort to reach the finish line first. A variety of variables can be changed to ensure the creation of four distinct levels of game difficulty. If a player collides with an NPC car during the race, their speed will reduce, and the car may be spun around the track. Players can be hit by NPC cars in any direction, and NPC cars can also cause collisions with the player car. Depending on the level of difficulty that was selected, players can collect and employ various boosts in order to help them succeed in the race. These boosts can be collected by the player from the track. To collect a boost, the player must manoeuvre their car around the track and drive through the boost to add it to their inventory. Once collected, weapon boosts (Rockets) can be fired used the L1 trigger on the PlayStation 3 controller, and protective boosts (Ghost Mode, Shields) can be applied using the R1 trigger. The functions of these collectable boosts are detailed below:

2.5.1.1 Rockets

Rockets are an offensive weapon that are available for collection during gameplay. A player can pick up rockets on the track by driving through the rocket pickup that sometimes appears (see Figure 17). Collecting a rocket pickup adds five rockets to the player's inventory. The maximum number of rockets that can be stored in the inventory at any time is 5. If a player shoots a rocket and hits another vehicle, the vehicle will stop and spin around the track, reducing their overall race performance. NPC vehicles can also collect and fire rockets, either at other NPCs or at the player car. Rockets always perform in the same way, regardless of which vehicle is targeted. An example of the use of Rockets can be seen below in Figure 17.

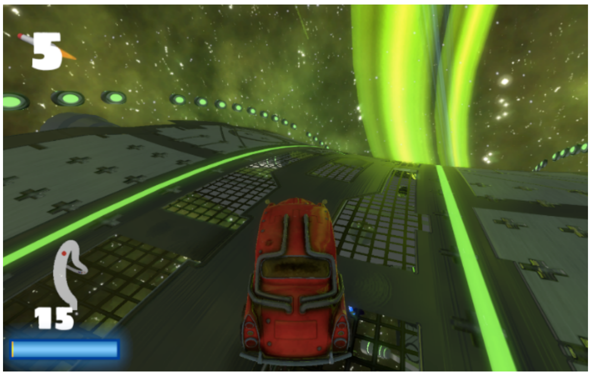


Figure 17 - Example of the use of Rockets in Space Ribbon

2.5.1.2 Shields

Shields are both an offensive and defensive mechanism that can be used during gameplay. A player can collect a Shield pickup by driving through the blue bubble pickup, which sometimes appears on the track. Collecting a Shield pickup adds one Shield to the player's inventory. The maximum number of Shields that can be stored in the inventory at any time is 1. With a Shield enabled, a grey, semi-transparent bubble appears around the player vehicle. When the Shield is activated, if a player drives into another vehicle then the vehicle will stop driving and flip into the air, before landing back on the track. It is possible that the attacked car will land upside down, meaning that the car has to be turned back around before they can continue to race. If another player drives into a car that has an activated Shield, the player will also be flipped into the air. Once a Shield is enabled, this pickup is active for 15 seconds before being automatically deactivated. An example of the use of Shield can be seen in Figure 18 below:

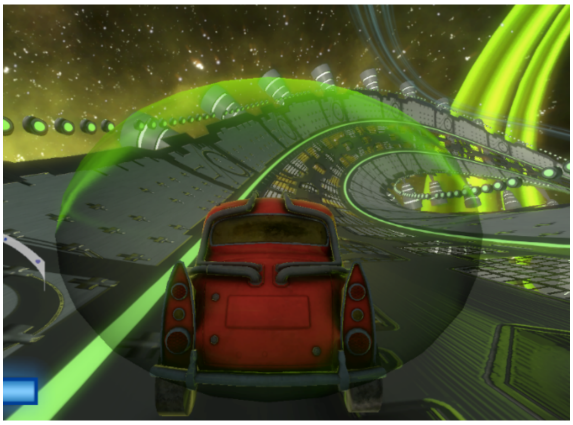


Figure 18 - Example of the use of Shield in Space Ribbon

2.5.1.3 Ghost Mode

Ghost Mode is a defensive mechanism that can be used during gameplay. A player can collect a Ghost Mode pickup by driving through the ghost shaped pickup that sometimes appears on the track. Colleting a Ghost Mode pickup adds one Ghost Mode shield to the player's inventory. The maximum number of any shields (either Ghost or Shield) that can be stored in the player's inventory at any time is 1. With Ghost Mode enabled, the players car becomes transparent and is able to move through NPC cars without causing a collision. The player can also drive through NPC vehicles, which are using Shields, without facing the usual repercussions. Once Ghost is enabled, this pickup is active for 15 seconds before being automatically deactivated. An example of Ghost can be seen below in Figure 19:

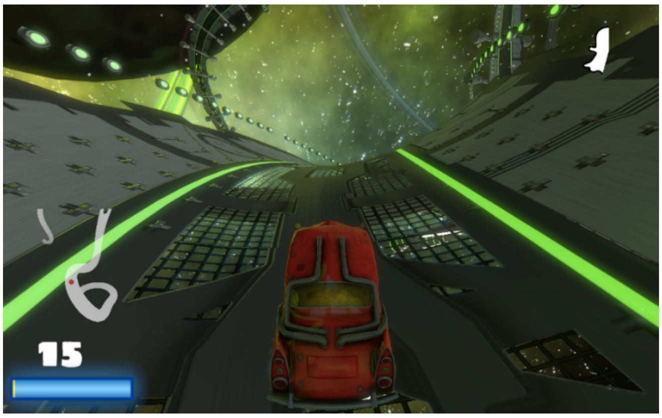


Figure 19 - Example of the use of Ghost Mode in Space Ribbon

2.5.2 Ribbon Rush

Ribbon Rush is an adapted version of *Space Ribbon*, which was used for Studies 2 and 3. This game did not follow the typical conventions of a racing game but was instead designed as a collision avoidance game. In this game, the player car was traveling against the flow of traffic, meaning that the goal of the game was to avoid a head-on collision with oncoming traffic. The goal of *Ribbon Rush* was not to win a race, but to maximise points over the duration of the game. Points were accumulated as the player moved along the track and were deducted every time the player collided with another vehicle. In *Ribbon Rush*, the goal of the game was to manoeuvre in such a way as to avoid as many collisions as possible. As such, there were no weapons or shields that the player could use. Although there was no weaponry, NPC vehicles did have the ability to swerve into the player car, which required the player to quickly manoeuvre or face the risk of a collision. The difficulty of *Ribbon Rush* was adapted simply by adjusting the speed of the game. For instance, an increase in game speed meant that both the player car and any NPC vehicles became faster, allowing the player less time to consider and carry out avoidance manoeuvres. Ribbon Rush did allow for the adaption of game difficulty via other means, such as road geometry and number of vehicles on the track, but it was determined that the most efficient and consistent method for difficulty adaption would be to alter just the speed of the game. Altering the speed of the game in turn altered the overall demand which was required of the player.

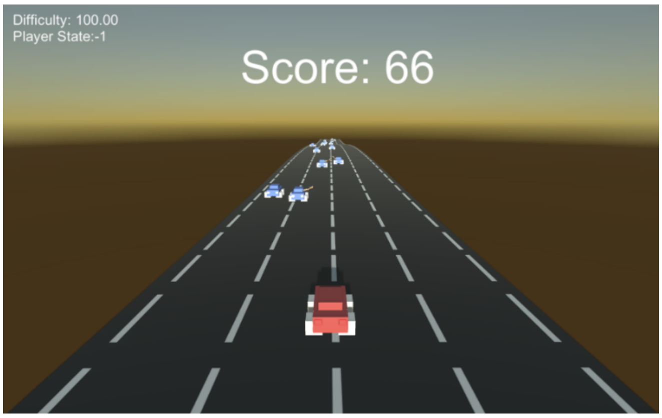


Figure 20 - Example of Ribbon Rush gameplay

2.6 Feature Selection and Data Classification

One purpose of the overall research project was to classify player response to game demand. Classification was carried out using multiple measures, both individually and in combination. The measures that were used for classification were fNIRS, ECG and game score. fNIRS data has previously been used to classify workload (Ayaz et al., 2013; Heger, Herff, Putze, Mutter, & Schultz, 2014; Herff et al., 2014). Our work concerns creating a real-time closed-loop system, with one overarching priority of a real-time system being that it is computationally efficient. Improving the computational efficiency of real-time classification can be achieved via defining a minimal set of critical features from the collected data (feature selection). The chosen features must be capable of differentiating between high and low game demand, or between pain and no pain. Classification can then be used to determine changes, which may be required to game demand. This section describes our approach to classification, which was used throughout all studies.

2.6.1 Feature Extraction and Selection

Feature extraction refers to the extraction of features from filtered data, which can provide a more detailed insight as to the performance or mental state of the player. Feature extraction can refer to statistical results (mean, range etc.) or results concerning a combination of different data points (connectivity measures). Each study in this research project used different feature extraction techniques, therefore the specific techniques will be discussed in more detail in the relevant chapters (Chapters 3, 4 and 6.). Feature selection is a vital precursor to machine learning, which is used to reduce the overall number of features within a dataset. As previously discussed, this step is important for improving computational efficiency during classification, whilst still maintaining maximum effectiveness. When feature selection is applied, the most relevant information from the dataset is maintained for classification.

2.6.1.1 RELIEFF Algorithm

Feature selection was undertaken using the RELIEFF algorithm (Kononenko, Šimec, & Robnik-Šikonja, 1997). RELIEFF uses a *k* nearest neighbour approach to weight the estimated quality of features. The *k* value is the value that determines the number of nearest neighbours that should be compared to each data point. This is done in order to determine the nearest values in the same class (hits) and the nearest values in a different class (misses). Each feature is weighted to estimate its quality, based on the amount of hits and misses. RELIEFF was performed over the datasets and the resulting weights were then ordered, from highest to lowest, and plotted on a graph (see Fig 21). The point in the graph where the "elbow" appears indicates the most relevant features and the feature set is cut down at this point. When features cease to provide relevant information for classification, they can be removed from the dataset to help prevent overfitting and reduce the overall classification time. Features which are removed from the dataset are deemed to be irrelevant to the classification – either because the provide no more information than features that have already been included, or because they don't follow an identifiable pattern between classes.

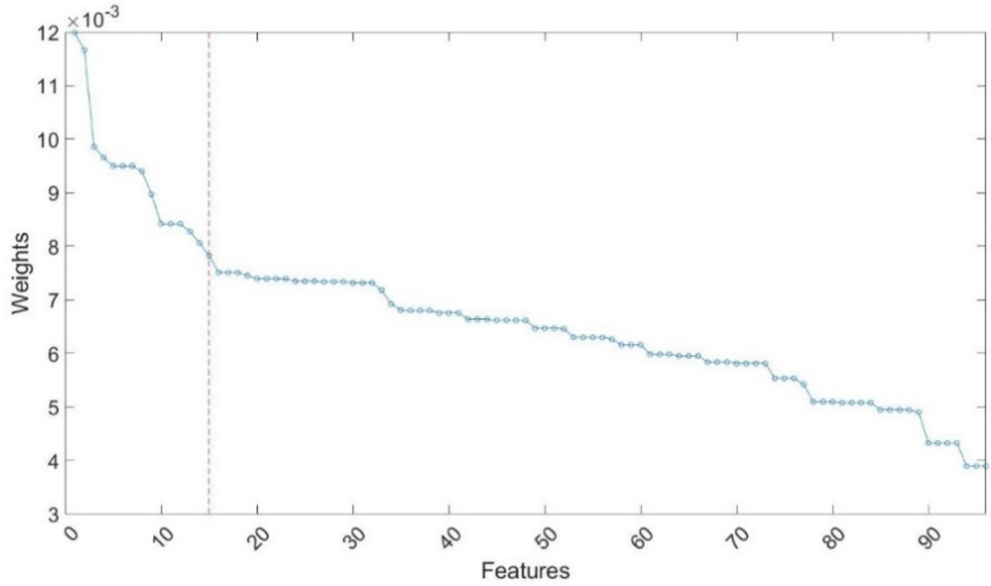


Figure 21 - Example of RELEIFF Algorithm Feature Selection

Data were also separated through each individual study in a variety of ways, including separating the data into frontal and central sites (Study 1) and separating the data into connectivity measures and average features (Study 2). Data separation techniques are discussed in more detail in the relevant chapters (Chapters 3 and 4).

2.6.1.2 Paired Sample T-Tests

Paired sample t-tests can be used to determine statistical significance between two variables gathered from the same participant under different conditions. In our case, we used paired sample t-tests to determine whether there was statistical significance between individual features. Where there is a statistical significance between pairs of features, these features are deemed to be relevant for classification, as they provide different data points, which can inform classification. Feature selection, via paired sample t-tests, enables us to identify the features that are relevant for classification, as the discarded features provide similar data to data that has already been included. This ensures that each feature that is chosen is providing a statistically significant influence to the dataset. Features which were determined to be statistically insignificant were removed from the dataset manually prior to classification.

2.6.2 Supervised Machine Learning

Machine Learning was used throughout this project to assess the accuracy of a variety of features for the classification of game difficulty. In the case of Studies 1 and 2, supervised machine learning classification was applied offline. However, in Study 3, machine learning techniques were implemented in real time in order to provide classifications for the adaptive game. The classification algorithms used throughout this project included both parametric and non-parametric algorithms, including Liner Discriminant Analysis (LDA) and Support Vector Machines (SVM). As both LDA and SVM are supervised learners, labelled training data is always supplied to the classifier for comparison against labelled, untrained test data. LDA and SVM techniques use a hyperplane to separate the data into two classes. A hyperplane can be seen as a line that is drawn between the plotted data and separates one class from another. An example of a hyperplane can be seen below in Figure 22[7].

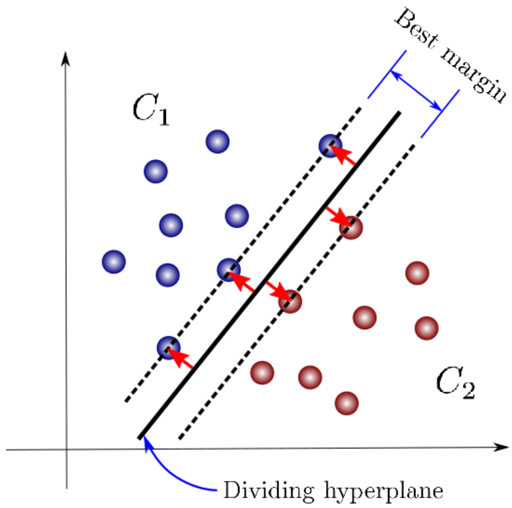


Figure 22 - An example of LDA hyperplane techniques

Although both SVM and LDA use a hyperplane to separate data, they apply different approaches to the determine the positioning of the hyperplane. Parametric classification algorithms (such as LDA) estimate the coefficients of the line that separates the data. LDA's are designed to separate each data point into one of two conditions, and a hyperplane is drawn that best separates all of the data into one of the two categories. This means that, if data is not evenly distributed, it becomes difficulty to draw a hyperplane that can discriminate between the two categories. This results in reduced accuracy, due to the discrepancy of which data falls into which categories. Nevertheless, the benefits of parametric algorithms are that they are simple, fast and do not require large amounts of training data. However, if the data that is being classified is not linear, then the algorithm can have a poor fit to the data and is best used in simple classification problems where it is expected that the data is evenly distributed.

In contrast, non-parametric algorithms (such as SVMs) do not make assumptions about the dataset, and therefore learn entirely from the training data. Non-parametric algorithms are more flexible and can perform better than parametric algorithms. However, they are more time intensive, require a larger data set for training purposes and are prone to overfitting, wherein assumptions are made about the data based on outliers rather than genuine data patterns. Furthermore, in contrast to the hyperplane techniques used by LDA's, SVM's draw a hyperplane between only the data points within two classes that fall closest to the group separation boundary. This means that SVM's can better categorise more randomly distributed data but rely only on a subset of data rather than the entire data set. This means in turn that deviations in data (due to potential spikes etc.) can be miscalculated as being inherent within one class. Therefore, SVMs can be unreliable in cases where the data is highly variable.

[7] https://towardsdatascience.com/support-vector-machines-for-classification-fc7c1565e3).

The purpose of using both LDAs and SVMs for classification was to determine which algorithm was best suited for the different data that was to be classified. There is an inherent bias-variance trade-off between LDAs and SVMs. LDAs are a restrictive technique, which have a high bias but introduce errors by making approximations of complex data with a simple model. SVMs have a high variance, which means that they are highly dependent on the training set, which can lead to a large variation in fit. By using both LDA and SVM techniques, we believe that we have identified two contrasting classification techniques which are both effective in the classification of fNIRS data (Hong et al. 2018) and are appropriate for real-time use. The final goal of this research project was to create a real-time system, so it was important that the time constrains of the chosen classification techniques were considered. Research and testing of ensemble methods was undertaken – ensemble learners combine the strongest elements of a chosen set of classification techniques in order to create a stronger classification. However, our work indicated that using an ensemble learner in a real-time situation would not be appropriate, as the classification could not be carried out in an appropriate timeframe. For this reason, we chose to maintain the use of both SVM and LDA classification thorough Studies 1 and 2, to enable us to make the most appropriate choice when it came to the real-time application of the classification techniques.

Classification carried out during this project was undertaken using RStudio Version 1.1.414 and the Machine Learning in R (MLR) package. The results were validated using *k*-fold cross validation, where $k = 10$. Performance measurements that were calculated included:

- *Accuracy* – an overall score of the performance. This score represents the percentage of the dataset which was correctly classified. Measuring the accuracy of the classification provides an overall indication of how affective the classifier was.
- *F_1 Score* – the weighted average of the precision and recall of the classifier. Precision and recall are both measures which provide an insight into where the classifier was successful or unsuccessful. Precision is calculated as the percentage of true positives from all positives that were estimated. For example, if the classifier predicted that 60 out of 100 cases should be classified as 'Hard', but there were only 50 true 'hard' cases, then the precision would be 83.3 percent. Precision is calculated by dividing the true positive amount by the true positive and false positive amount and multiplying the product by 100. Equation *n* below illustrates this example as a calculation:

$$(50 \div 60) \times 100 = 83.3 \qquad (n)$$

Recall is calculated as the percentage of positive classifications that were made compared to the number of true positives. Recall gives a measure of how effective the classification is at finding true classifications. Referring back to the earlier example, if the classifier predicted 60 'hard' classifications out of 100 cases (although only 50 cases were true positives) then there are 0 false negatives – and therefore the recall rate is 100%. Using a combination of precision and recall to calculate the F1 score allows for the determination of accuracy for determining both true positives and true negatives.

- *Balanced Error Rate (BER)* – an average of classification errors that occur in each class. BER is used to calculate the mean misclassification rate of each class that is available for classification.

2.7 Summary

This has presented information relating to equipment, data collection protocols, data collection and processing techniques and data classification that were consistently utilised throughout each study undertaken as part of this research project. Chapter 2 serves to introduce the reader to common themes that will be further elaborated upon throughout the remainder of the research.

Chapter 3 - Study One: Examining the Experience of Pain and Distraction in the Prefrontal and Somatosensory Cortex

3.1 Introduction

This chapter will focus on the collection and classification of fNIRS data relating both to painful experiences and a gaming distraction technique. The materials and methods used during data collection will be detailed, and the techniques used to process and classify the data is be discussed. This chapter will end with a presentation of results and subsequent discussion, aiming to assess the validity both of measuring pain and distraction with fNIRS equipment, and the validity of the chosen classification measures.

Distraction techniques are effective in reducing the perception of pain due to selective attention. Pain is considered to be a threatening form of stimulation that can interrupt attention to other stimuli in the environment (Eccleston & Crombez, 1999). However, directing attention away from painful stimuli through goal-orientated tasks can modulate this interruptive function of pain. For example, fMRI has been used to capture neurovascular activation during painful stimulation with and without distraction (Bantick et al., 2002). The authors reported a reduction of activation in the pain matrix (thalamus, insula, anterior cingulate cortex) in the presence of a distraction (compared to the control condition). This evidence suggests that analgesic benefits are derived from distracting stimuli working directly on the relevant brain areas, specifically the anterior cingulate cortex in this case (Bantick et al., 2002; Legrain et al., 2009).

In order to maximise the effectiveness of a distraction approach, the distracting stimuli or task should require a high level of cognitive effort in order to draw attention from the painful stimuli (Johnson, 2005). There are a wide variety of techniques that may be used to distract from pain, such as watching the television or reading a book, however computer games have been proven to be the most effective approach (Bantick et al., 2002; Johnson, 2005; Legrain et al., 2009). The act of playing a computer game functions as an active distraction, which requires effortful pursuit of game-related goals and focused attention, whereas watching television is relatively passive in comparison. While the psychological demands of a computer game create an immersive experience that actively distracts from pain stimulation, achieving this immersive experience may not be straightforward (Hoffman, 2004). Fairclough et al. (S. H. Fairclough, Gilleade, Ewing, & Roberts, 2013a) (2013) noted that most games are aimed at an 'ideal player' who does not actually exist. Some players will have more or less experience than said 'ideal player' and will therefore find a standard game to be either too easy or too difficult. Therefore, it is important that the demands or difficulty of the game are optimized to engage the individual, according to his or her capabilities, in order to distract from pain. Csikszentmihalyi (Csikszentmihalyi, 1975) coined the term 'flow' in reference to the psychological state experienced by a person when they are totally involved with a task, which is true also of total immersion (Csikszentmihalyi, 1975; Jennett et al., 2008).

For an active distraction task to provide an analgesic effect, the mitigating factor is task demand. The cognitive demand of the task requires focused attention, which reduces the amount of attention that is available to attend to alternative stimuli, which, in this case, is pain. However, task demand is a personal experience. For example, a novice gamer may experience a high level of engagement towards a game that an experienced gamer considers to be too easy, and therefore boring. This represents an interaction between objective game demand and the skill of the player. Therefore, for a game to effectively distract a player from pain, it must maximally engage the player, which requires an adjustment of the game mechanism based on the *individual* skill level. If a game is too easy, then the player will lose motivation to continue to play. In the same way, if the game is too hard, then the player will become disengaged (as they become aware that they cannot win) and give up playing. It is imperative for a distraction task to retain the players attention in order for said player to experience analgesic effects.

In this study, the effects of game demand and experimental pain have been measured using behavioral, subjective and neurovascular activation measures from frontal and central areas of the brain. This study will develop a classification technique that can be used to assess the level of game demand, using implicit measures of brain activation. The purpose of the study is to explore the relationship between game demand and pain tolerance. It is hypothesized that high demand will increase pain tolerance compared to low game demand. It is also assumed that the highest level of game demand will be less effective as a distraction from pain because success becomes highly unlikely and therefore, we expect attention to reduce. The central goal of this study is to assess the ability of fNIRS data to differentiate between high and low game demand, to inform the viability of using fNIRS to control a neuroadaptive game. We also wished to explore the relationship between game demand and pain tolerance, to determine whether a more immersive gaming experience would provide a higher level of pain tolerance.

The fNIRS montage used in Study 1 was designed to collect data from the prefrontal cortex and the somatosensory cortex as independent brain areas. The prefrontal cortex is thought to be involved in decision making and executive function – which includes the ability to determine and differentiate between options relating to the outcome of a defined goal and the prediction of outcomes. The prefrontal cortex was targeted as we expected to see activity relating to gameplay decisions and outcomes occurring in this area of the brain. The somatosensory cortex is part of the somatosensory system, which plays an important role in the perception of pain. We targeted the somatosensory cortex with the intention of collecting data which could relate to the pain that the participant would experience. Our hypothesis was that we would see heightened activity in the somatosensory areas when the painful stimulus was introduced.

3.2 Materials and Methods

3.2.1 Design of study

The study was designed as a within-participants design, wherein all participants were exposed to 4 levels of game demand (easy, medium, hard and impossible) under two conditions – with or without experimental pain induced via a cold-pressor test (CPT).

3.2.2 Participants

Data was collected from 20 participants, of whom 6 were female. Participants were aged between 19 and 29 *(M = 22.75 SD = 3.23).* Participants were excluded from participating if they were pregnant, or if they had a history of cardiovascular disease, fainting, seizures, chronic or current pain, Reynaud's disease or diabetes. Participants with fractures or open cuts or sores on the feet or calves were also excluded. Participants were required to confirm that they were not currently taking any medication, with the exception of the contraceptive pill. A full review of the ethics of the experiment was undertaken, and approval was granted by the Liverpool John Moores University Research Ethics Committee. All participants were briefed before their experimental session and were provided with a detailed Participant Information Sheet prior to taking part. Full written consent was provided by each participant involved in this study.

3.2.3 Apparatus

fNIRS data was collected using an Artinis OxyMon Mk III device, which measured neurovascular activity at the cortex (see Figure 23). A 2x4 cross-channel configuration was used, with a total of 2 sources and 8 detectors. 4 channels were situated at the prefrontal cortex, between Fz and: F1, AFZ, F2 and FCz, and the remaining 4 channels were situated at the somatosensory cortex, between CPZ and: CP1, Cz, CP2 and Pz. This optode layout is illustrated in Figure 23. It is important to note that data was collected from the areas between the transmitter and receiver, and not from the receiver itself. Source-detector separation was 3cm and source optodes emitted light at 847nm and 761nm wavelengths. The device was configured to record optical density data at a sampling rate of 10Hz. Data was recorded using the Oxysoft data recording software (Artinis).

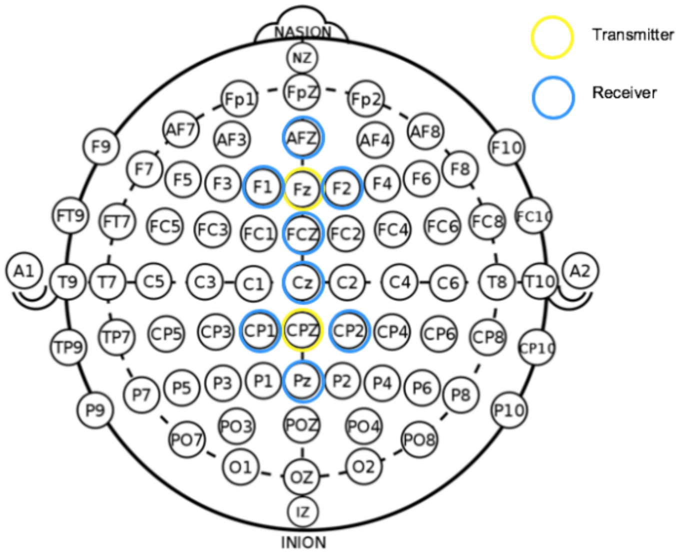


Figure 23 - fNIRS Optode Placement. Two light sources (central), and 8 detectors surrounding the light source.

To counteract the effects of potential head movement noise, it was decided that an accelerometer would be used to record head movement data, so it could be subsequently filtered out of the signal. The accelerometer that was used was a Shimmer3™ inertial measurement unit (IMU). This device had a sampling rate of 512 Hz. The Shimmer unit was worn around the participants head (over the fNIRS cap) on an elasticated band.

3.2.4 Cold Pressor

A bespoke device was created (Dancer Designs) consisting of two water tanks, pumps and thermostat. One tank allowed the participant to place a limb (in this case, foot) in cold water that was regulated via negative control loop to maintain a fixed temperature of 2 degrees. The other tank fed this immersion tank to ensure that water temperature was sustained. Refer to Chapter 2 for full details of the Cold Pressor device and test.

3.2.5 Game

Four levels of game difficulty were created in Space Ribbon (see Chapter 2 for details): Easy, Medium, Hard and Impossible. We expected that the Easy level would be too boring for the majority of participants, and they would easily achieve and maintain the 1st place position. In contrast, we expected that the Impossible level would be incredibly difficult, and that most, if not all, participants would be unable to achieve a 1st place position in this race. It was expected that the two middle levels, Medium and Hard, would provide a more immersive experience for the participants, as they were designed to be difficult but ultimately achievable. A pilot test of five participants was carried out to ensure that participants considered each level of game difficulty to be different from the other. Game difficulty level was manipulated by increasing the number of Non-Playable Characters (NPCs), increasing the artificial intelligence (AI) of the NPCs, and increasing the speed of the race overall. Increasing the AI of the

NPCs meant that they were more likely to pick up shields, rockets and ghost mode activators, which gave them an advantage over the player, and improved their position on the track and enabled them to pick up more speed. Table 1 below indicates the differences between the four levels of game difficulty, and the variables that were altered to establish these difficulty levels. For a detailed description of each of the individual variables, refer to Chapter 2.

Table 1 - Game variables which were altered to define four individual difficulty levels

Variables	Easy	Medium	Hard	Impossible
NPCs	3	5	9	15
AI Difficulty	1	2.5	3.5	5
Race Speed	1	1	3	5
Manoeuvre Speed	1	1.2	1.4	1.6
Shields	Off	On	On	On
Rockets	On	On	On	On
Ghost	On	On	On	Off

3.2.6 Subjective Questionnaires

As described in Chapter 2, a number of questionnaires were used to measure pain perception, mental workload and overall motivation. The NASA Task Load Index (TLX) was used as a measure of the level of subjective mental workload experienced by participants. The Motivation scale was used as a measure of intrinsic motivation. The Immersive Experience Questionnaire (IEQ) was used to measure perceptions of immersion.

The IEQ, TLX and Motivation scales were completed after each game condition. These questionnaires were primarily used to determine the differences in game levels, and how these differences may have affected a participant's immersion levels and pain tolerance/perception.

3.3 fNIRS Data Pre-Processing

3.3.1 Data Filtering

During pre-processing, a variety of filters and algorithms were applied to the raw data (see Figure 24). Filtering of fNIRS data is required to ensure that the signal is free from noise which could obscure the clarity of the genuine brain activity, which is recorded in the signal. Filtering the fNIRS data ensures that analysis and classification of the data are as accurate as possible, and that the data are not reflecting false positives or negatives, which can lead to inaccurate classification and analysis.

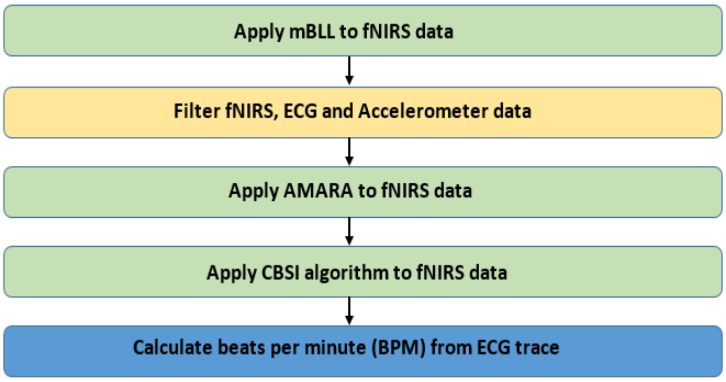


Figure 24 - Data pre-processing pipeline

In order to determine activation in the target cortices, the Optical Density data that was collected via the fNIRS cap were converted into readings of oxygenated hemoglobin (HbO) and deoxygenated hemoglobin (HHb), using the modified Beer Lambert Law (mBLL) (Baker et al., 2014). Following this conversion, the fNIRS data was filtered using a 6^{th} Order Chebyshev filter, with passband edge frequencies of 0.5 and 0.1 Hz respectively, for high and low pass filtering. These filters were applied to reduce noise within the signal that related to heart rate and respiration as well as Mayer waves, which are expected to occur within the fNIRS signal due to changes in arterial pressure (Naseer & Hong, 2015). It is important to remove these features from the fNIRS signal to ensure that the signal reflects a true indication of neurovascular response, rather than responses that relate to physiological changes, which can occur at the same frequencies (Julien, 2006). The same filters were also applied to the electrocardiogram (ECG) and accelerometer data to reduce noise within these signals.

Following high and low pass filtering, both fNIRS and accelerometer data were processed using the Acceleration-Based Movement Artefact Reduction Algorithm (AMARA) (Metz et al., 2015).

It has been reported that head movement artefacts can also cause a positive correlation between HbO and HHb, where these signals would usually be negatively correlated. Therefore, if the signals show that there isn't a strong negative correlation between HbO and HHb, it could be an indication that there is noise contained within the signal, which is having an adverse effect. To counteract any movement artefacts that could still be having a negative effect on the data even after the application of the AMARA, a Correlation Based Signal Improvement (CBSI) algorithm (Cui et al., 2010a), which corrects the correlation of the HbO and HHb signals, was applied to the data.

In order to determine the heart rate for classification, equation 5 was used to calculate beats per minute (BPM) from the ECG signal. Calculating BPM provides a numerical measure of heart rate data which can be used to determine how, if at all, different conditions will affect a participants HR.

$$BPM = \frac{6000}{(rPeakECG2 - rPeakECG1) * 1000} \tag{5}$$

This process involved calculating the inter-beat interval (IBI), which corresponds to the time between consecutive R waves in an ECG signal. This was undertaken by finding all of the R peaks in the signal and then finding the difference in time between two successive peaks. Once this time has been calculated, the time between the peaks is converted into milliseconds. This data is then converted into beats per minute (BPM) by dividing by 60,000 (the number of milliseconds in a minute) by the time (in milliseconds) between each peak, as in equation *n*. This provides an accurate representation of the BPM each time the heart beats.

3.3.2 Feature Extraction

Following pre-processing, Total Hemoglobin (HbT) was calculated, which is created through addition of the HbO and HHb signals to provide a signal that details the total cortical activity, as in equation 6.

$$HbT = HbO + Hbb \tag{6}$$

The fNIRS data were then separated into 8 second windows. This was due to the hemodynamic Delay that is present in fNIRS data, where the response to the onset of stimuli has a delay of several seconds after the stimuli have been introduced, before changes in the signal reflect this (Mayer et al., 2014).

As the 8 channels were split evenly between the Prefrontal and Somatosensory cortex, the decision was made to investigate brain connectivity. Connectivity measures identify correlations between different areas of the brain that may be working in conjunction with each other. The underlying communication dynamic can be observed via oscillations within the brain between these areas. If these oscillations are examined, a greater picture of the functional architecture of the brain may be observed (Bastos & Schoffelen, 2016; Verdière, Roy, & Dehais, 2018). By creating connectivity features from the fNIRS data that are collected, we expect that we may achieve a greater classification accuracy than using standard features (e.g. Peak, Variance etc.) alone. Connectivity measures were calculated for HbO and tHb over 8 second epochs. At this point, HHb data was excluded from feature extraction, due to the function of the CBSI filter. It was determined that, as HHb data was now entirely negatively correlated from HbO data, there would be no further useful activity that could be determined from the HHb signal. The nature of connectivity measures means that features were created for each possible pair of channels. As such, 8 channels provided 28 possible pairs of channels. For each of these 28 channels and each channel's respective 2 features, there were 56 types of data from which connectivity measures were created. For each of these 56 instances, 5 connectivity measures were calculated:

- *Covariance* (Equation 7) refers to a measure of joint variability of two independent signals. Covariance can either be positive or negative, and the sign of the covariance shows the tendency in the linear relationship between two variables.

$$COV(x, y) = E\big(x - E(x)\big) * E\big(y - E(y)\big) \tag{7}$$

- *Pearson's Correlation Coefficient* (Equation 8) is the covariance for two signals, normalized by their standard deviation. This correlation coefficient can be either linearly positive, linearly negative, or show no correlation.

$$Pearson(x, y) = \frac{COV(x, y)}{std(x) * std(y)} \tag{8}$$

- *Spearman's Correlation Coefficient* (Equation 9) is a nonparametric measure of rank correlation, which is the statistical dependence between the rankings of two variables.

$$Spearman\ (x, y) = \frac{COV(rgx, rgy)}{std(rgx) * std(rgy)} \tag{9}$$

- *Spectral Coherence* can be defined as "*the absolute squared value of the cross-spectral density of two signals, normalized by the product of their auto-spectral density*" (Verdière et al., 2018).

$$Cxy(f) = \frac{Gxy(f)2}{Gxy(f) * Gyy(f)} \tag{10}$$

- *Wavelet Coherence* (Equation 11) can be seen as "*a localized correlation coefficient in time frequency space*" (Grinsted, Moore, & Jevrejeva, 2004). The values obtained were averaged for frequencies between 0.3125 Hz and 0.08 Hz, as recommended by the fNIRS literature (Cui, Bryant, & Reiss, 2012).

$$R_n^2(s) = \frac{|S(s^{-1}W_n^{xy}(s)|^2}{S(s^{-1}|W_n^x(s)|^2)S\left(s^{-1}\left|W_n^y(s)\right|^2\right)} \tag{11}$$

Calculating these features for each of the 56 instances of data provided us with 280 features. In total there were 8,400 instances of data for the connectivity measure analysis (280 (features) x 30 (epochs)).

Feature selection is a process whereby a dataset is reduced in size without reducing the efficacy of the data. Feature selection is carried out for a variety of reasons. The primary reason for the application of feature selection is to reduce the training time of the classification algorithm. A smaller feature set will provide a more time-efficient classification result than a larger feature set. Feature selection can also improve the accuracy of a machine learning model and reduce overfitting. Improving the accuracy of the model can be achieved by removing features which are not relevant to the dataset for the chosen classification. Reducing overfitting requires reducing the complexity of the model, to prevent the training data from being too heavily informed by patterns in the data which are not relevant for classification. Classifications which have been vulnerable to overfitting will appear to be accurate in a training/test set comparison but will not maintain accuracy as new data is introduce into the model. Feature selection for Study 1 was carried out by using the results of an independent t-test.

Classification of the feature-selected dataset was carried out using supervised machine learning classification. Supervised classification techniques were chosen due to their appropriateness for real-time classification. The aim of Study 1 was to explore the potential for a real-time adaptive game, so suitable real-time algorithms were chosen for this purpose. Supervised learners are more appropriate for real-time classification because the training data is already labelled prior to classification – which means that subject independent adaptation could be carried out more quickly than if unsupervised learning was used.

3.4 Procedure

Before the experiment began, each participant was presented with an information sheet that explained the procedures that would be followed, and what would be required of them during the study. The participant then signed a consent form, and was familiarized with both Space Ribbon and the CPT. Following familiarization, the participant was then instructed on how to place the Bioharness on their chest and was given privacy whilst they attached the equipment to their body. The bioharness signal was checked by the experimenter to ensure correct placement of the equipment. Then, the fNIRS cap was placed on the participants head, and the accelerometer band was placed over the fNIRS cap. The participant was now ready to begin the experiment.

Figure 25 illustrates the six steps of the experimental protocol that were undertaken. To ensure that the results collected during each step of this study were independent, a 90 second baseline period was established between each condition. The entire protocol was repeated a total of four times, to enable participants to play each of the four levels of game difficulty (i.e. easy, medium, hard and impossible).

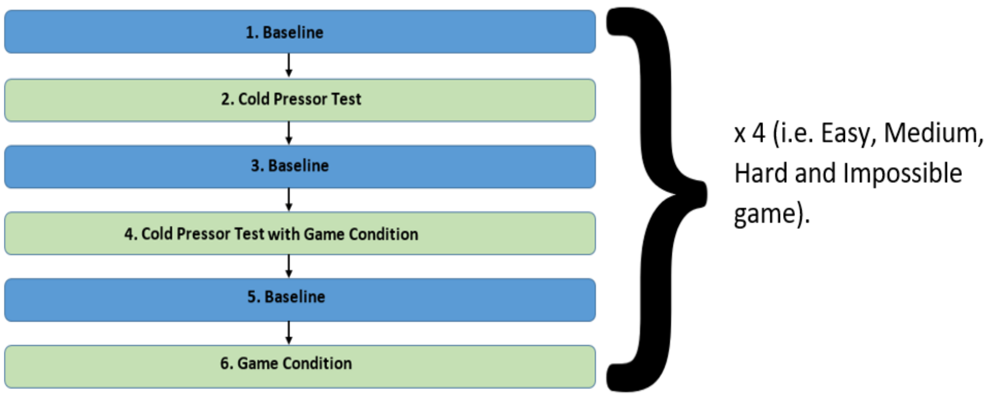


Figure 25 - An illustration of each step of the data collection protocol.

Throughout the course of the experiment, participants played all four levels of the game in a randomized order, to ensure that the results collected were not influenced by the participant gaining more experience with the game on Easy and Medium levels before they then played Hard and Impossible levels. This ensures that the participant's response to each level of the game is independent of their level of skill or their familiarity with the game. In total, 690,028 instances of fNIRS data, 16,763,628 instances of heart rate data and 26,177,699 instances of accelerometer data were collected, for a total of 43,631,355 instances of data overall. These six steps were repeated for each level of game difficulty. During each set of six steps, the save level of game difficulty was played for both step 4 and step 6.

3.5 Results

3.5.1 Behavioural and Subjective Measures

A one-factor ANOVA was conducted on the behavioral pain data (i.e. time with limb immersed in water) in order to understand how game difficulty influenced pain tolerance. However, three participants were excluded from the analysis because they kept their limb in the water for the maximum duration in all game difficulty conditions. The analysis of remaining 17 participants revealed that the main effect for game difficulty fell outside of statistical significance [$F(3,14) = 2.59$, $p=.09$]. Inspection of the mean values (Figure 26) illustrate that pain tolerance was lower during the Easy level of game difficulty compared to all others.

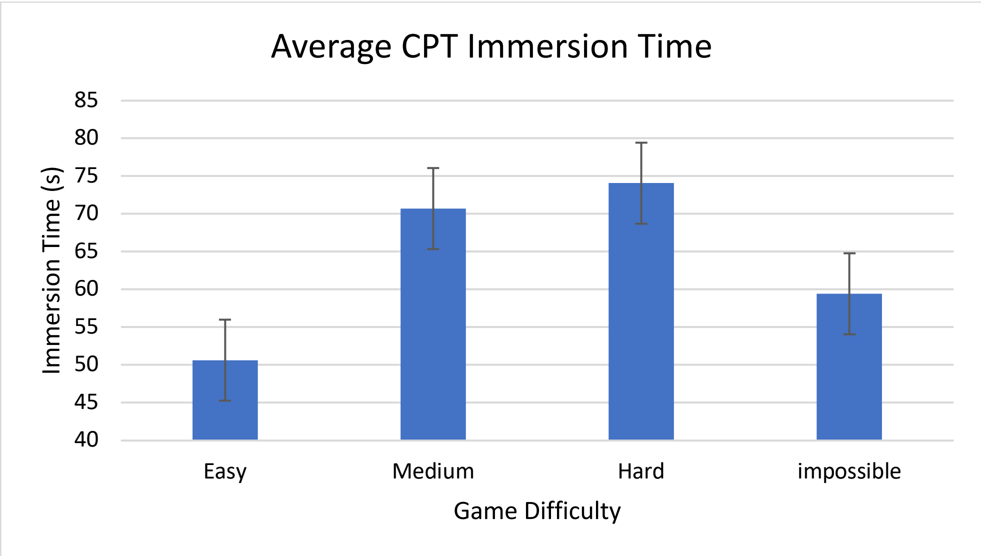


Figure 26 - Mean times that limb was immersed in cold water during CPT (N=17)

Participants completed three subjective questionnaires (TLX, Motivation, IEQ) after completing each game in all conditions. Figures 27-29 below illustrates the average results from each subjective questionnaire (TLX, Motivation and IEQ).

Figure 27 indicates the expected pattern of subjective mental workload across each level of demand. A one-way ANOVA was performed to test for differences in subjective workload across all four levels of demand. The analysis revealed a significant main effect for demand [$F(3,15) = 7.15$, $p<.01$, $\eta^2=0.59$]. Post-hoc tests revealed that subjective workload was significantly lower in the easy condition compared to all other levels of game demand ($p<.01$).

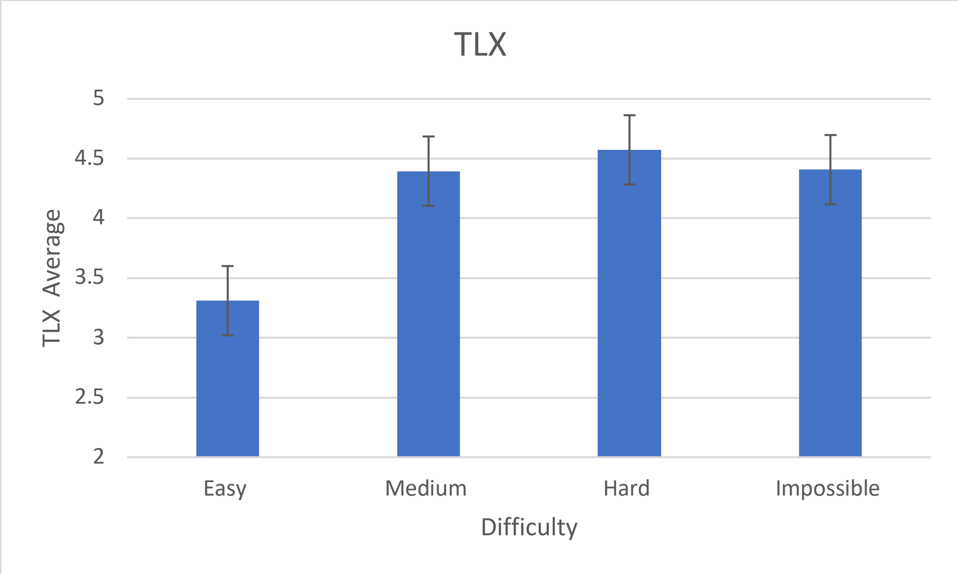


Figure 27 - Mean TLX Results

Figure 28 illustrates the results of the Motivation questionnaire. A one-way ANOVA was performed to test for differences in motivation across all four levels of demand, but none of the data trends observed in Figure 27 were statistically significant [$F(3,17) = 1.23$, $p = .33$, $\eta^2 = .178$].

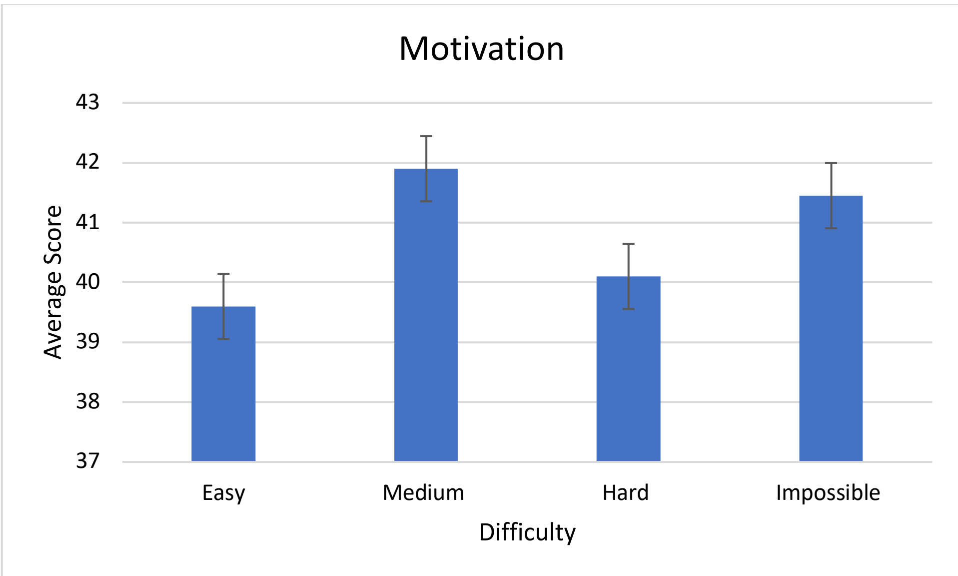


Figure 28 - Mean Motivation Results

Figure 29 illustrates the results of the IEQ. As expected, participants found the Easy level of demand to be less subjectively immersive than higher game demand. However, the ANOVA revealed main effect for demand of marginal statistical significance [$F(3,17) = 2.84$, $p<.07$, $\eta^2=0.33$]. Post-hoc tests revealed immersion was significantly lower in the Easy condition compared to all other levels of game demand ($p<.03$).

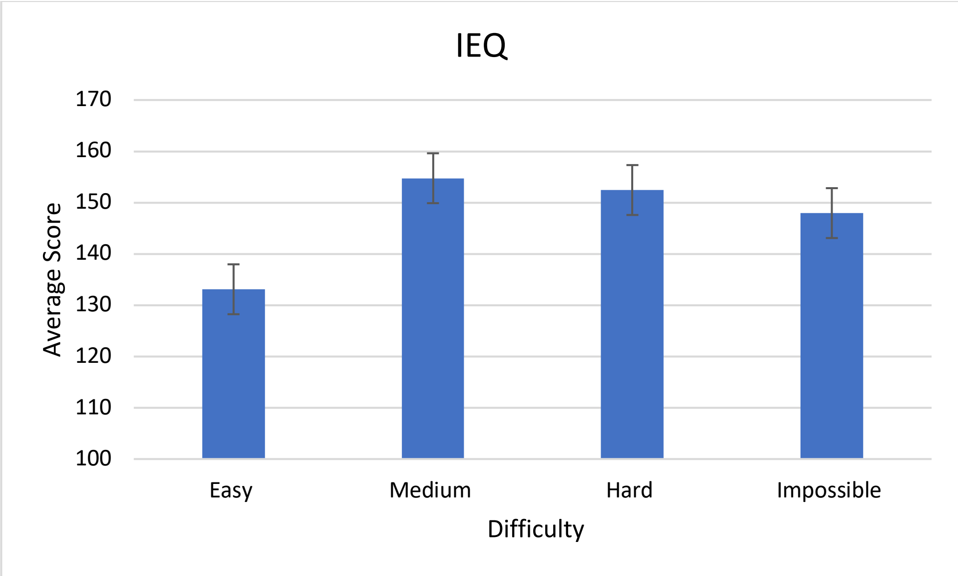


Figure 29 - Mean IEQ Results

3.5.2 Data Classification

A series of binary classifications were performed in order to determine the level of accuracy with which the different levels of game demand could be classified. These classifications were performed over two datasets. The first dataset was comprised of HR data, the second dataset was comprised of fNIRS data. The largest difference in pain tolerance was found between the Easy and Hard game conditions, which indicates that the more significant difference between game difficulty levels would be found between these conditions. For this reason, it was determined that binary classification should be carried out between Easy and Hard levels. In order to determine how accurately the levels of game difficulty can be classified, Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) classifications were undertaken. The accuracy, balanced error rate (BER) and F1 scores of these classifications will be reported. In order to determine which dataset (either HR or fNIRS) provides a more accurate classification, the same classification will be performed over both sets of data individually. To select the fNIRS features that should be used for classification, paired t-tests will be conducted. Only features that significantly differentiate one level from another at p=0.01 will be retained for classification.

3.5.2.1 Feature Selection

Feature selection was carried out for via T-Tests, which were conducted to find out the significance levels of each individual feature. Only the features that had a significance level of p=0.01 were selected for classification. Table 2 below indicates the features that were chosen for classification from this feature set using independent T-Tests.

Table 2 - Relevant fNIRS features selected for classification determined by independent T Tests

tHb	HbO
2_Skew	2_Average
4_Peak	1_AUC
3_Variance	2_AUC
4_Variance	3_AUC
1x2_Pearsons	1x2_Pearsons
1x3_Pearsons	2x3_Pearsons
2x3_Pearsons	1x3_Pearsons
2x4_Pearsons	1x5_Pearsons
3x4_Pearsons	2x6_Pearsons
1x5_Pearsons	3x7_Pearsons
2x6_Pearsons	
4x5_Pearsons	
1x7_Pearsons	
3x5_Pearsons	
3x8_Pearsons	
4x8_Pearsons	
2x8_Pearsons	
7x8_Pearsons	

This table denotes the channels (or pairs of channels) that were selected. Skew refers to Skewness, AUC refers to Area Under the Curve, and Pearson's refers to Pearson's Correlation Coefficient. See Figure 23 in the Method section for channel number and placement information.

3.5.2.2 Comparative Classification

Classification was carried out for this study in two stages. The first stage was carried out to establish the ability of HR data to classify game demand. In order to compare the efficacy of fNIRS data to more commonly used data, classification was also carried out using fNIRS data. These classifications were preformed to establish whether fNIRS data is necessary to provide an accurate classification score, or if the same classification rates can be achieved with a measure of autonomic nervous system activity, which is both easier to collect and more comfortable for the participant.

For this analysis, LDA and SVM were carried out over all four levels of game difficulty. On the basis of the patterns observed in the subjective data (TLX and IEQ) it was decided to classify between the Easy and Hard levels of difficulty. This decision was taken because the goal of the classifier was to differentiate those levels of game difficulty that had an observable impact on pain tolerance. Game difficulty was used to label data in the classification training sets. For example, all data that had been collected during the easy game was labeled easy. This convention was followed for all levels of difficulty.

During classification, data were cross-validated to detect and prevent overfitting within the machine learning process. In this case, k-fold cross-validation was used. K-fold cross validation splits the available training data into *k* number of subsets (folds). The model is then trained on *k-1* of the subsets of data. The remaining fold is used for classification evaluation. The *k*-fold cross-validation technique is repeated *k* times, with a different fold reserved for evaluation on each pass. For our research, *k* was equal to 10. Classification was evaluated using three measures – Accuracy, Balanced Error Rate (BER) and F1 score. Each of these measures provides insight into how the classification algorithm has performed, and how reliable the classification method can be considered.

Accuracy provides an overall score of the accuracy of the classification. As the classifications performed in this study were Binary (only two possible classification outcomes), there are four potential classification outcomes; True Positive (the classification returned a correct positive result), True Negative (the classification returned a correct negative result), False Positive (the classification returned a positive result which should have been negative) and False Negative (the classification returned a negative result which should have been positive). Accuracy can therefore be classified by dividing the number of True Positives and True Negatives (i.e. correct results) buy the sum of True Positives, True Negatives, False Positives and False Negatives. This equation is depicted below in Equation *1*:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Although accuracy can provide an insight into how a classifier is performing, it does not provide insight into how many cases are missed – for example, we may be able to see that a classification is accurate 80% of the time, but we cannot see how that is split between two results (Easy and Hard) and therefore cannot determine if the model is classifing disproportionally over one class.

BER provides insight into the average amount of errors (incorrect classifications) in each class. An example of this would be data which were classified as 'Hard; when the data was in fact collected during the easy level of the game. In the case of a binary classification (where there are only two possible classes), BER can be expressed as such:

$$BER = 0.5 \times \frac{FP}{TN + FP} + \frac{FN}{FN + TP}$$

Calculating BER alongside accuracy allows for a clearer understanding of how the accuracy is affected by misclassifications in each class and provides a secondary insight into the performance of the classifier.

Finally, F1 score was calculated. The F1 score provides a harmonic mean (the mean of two independent means) which provides insight into the mean of multiple values. In this case, the two values which are used to calculate the harmonic mean are Precision and Recall. Precision is the number of correctly identified positive results, divided by the number of all positive results (including results which may have been misclassified). Recall is the number of correctly identified positive results divided by the number of all samples which should have returned a positive result. Using an F1 score alongside other measures (such as Accuracy and BER) provides a more robust evaluation of the performance of a classification method. The F1 score can be calculated as in Equation *2,* where P = precision and R = recall*:*

$$F1 = \frac{2}{R-1+P-2} = 2\,.\frac{P\,.R}{P+R} = \frac{TP}{TP + \frac{1}{2}\,(FP + FN)}$$

Overall, we believe that the three classification evaluation methods used (Accuracy, BER and F1) provide a reliable insight into the viability of the classification method. It is expected that each of the evaluation techniques will provide differing results, as each technique is evaluating a different aspect of the overall classification. In the case of a large variance between Accuracy and F1 score for the same classification, it can be inferred that, although the model is making accurate assumptions, there are still many false positives or false negatives. Accuracy only takes into account the amount of correctly predicted classifications compared to all possible classifications and is not a valid evaluation method for unsymmetric datasets, where there is variance between the number of false positives and false negatives. If a classification is reporting a high accuracy but a low F1 score, we can determine that this classification is not performing well, and therefore changes should be made to the classification or the dataset to improve the balance of precision and recall.

Although multi-class classification could have been used throughout this study to classify between either of the four possible conditions (Easy, Medium, Hard and Impossible), multiclass classifications are typically both less accurate and more time-intensive than binary classification methods. As the overall aim of this work is to create a real-time classification algorithm for adaptation, it is important that the classification is as accurate as possible and can be performed as quickly as possible. Through testing, it was determined that multiclass classification would not be suitable for a real-time application, leading to the determination that binary classification would be the most appropriate method.

The results of this classification can be seen in Table 3.

Table 3 - Heart Rate Classification Results

Condition	Performance Measure	LDA	SVM
Easy v Hard	Accuracy	54.2%	**55.4%**
	BER	49.5%	**47.6%**
	F1	20.9%	**25.3%**

The results in Table 3 indicate that classification accuracy was poor for HR data. The achieved classification rates would not be appropriate for an accurate classification of game difficulty. Accuracies of approximately 55% were achieved via both the SVM and LDA classification, with BERs of 49% and 47% respectively. These results indicate that an accurate classification of game demand was only made for slightly more than half of the data points. This cannot be considered an accurate classification, as it is only slightly higher than a random decision.

3.5.2.3 fNIRS Features Classification

A second binary classification was performed on fNIRS data only, to explore whether fNIRS data could provide a more accurate level of classification compared to HR data. The resulting classification from the selected fNIRS features can be seen below in Table 4.

Table 4 - Classification of fNIRS Data

Condition	Performance Measure	LDA	SVM
Easy v Hard	Accuracy	81%	**86.4%**
	BER	18.8%	**13.6%**
	F1	80.2%	**86.4%**

As can be seen in Table 4, classification was much more accurate when fNIRS data were used. Both an LDA and SVM learner were able to more accurately classify the level of game difficulty (when taking into account accuracy, BER and F1) than HR data. The classification accuracy of the SVM when using fNIRS features was 86.4%. This is an improvement of 31% on the HR classification. These differences can also be observed across the remaining classification measures. There is a decrease of 34% in the BER for fNIRS classification, compared to HR classification, and an improvement of 61.1% for the F1 score. Whereas the results of HR classification are similar between both the SVM and the LDA classification (difference of 1.2% for accuracy, 1.9% for BER and 4.4% for F1 in favour of SVM), there is a clearer distinction in favour of SVM when considering the fNIRS data (improvement of 5.4% for accuracy, 5.2% for BER and 6.2%). The results gathered from both HR and fNIRS data classification indicate that fNIRS provides a far more accurate classification rate, and that an SVM classification can provide the most accurate results overall. These results are a vast improvement on the previous results and show promising accuracy when considering the future development of the real-time application.

3.6 Discussion

We expected that increased game demand would lead to increased attention to the game, which would in turn lead to an increase in pain tolerance. We also expected that we would find increased activation in the cortex as a response to high game demand. The rationale for this study was that the implicit measurement of demand experience by the player could be measured via fNIRS, which would provide and index for increased game demand and, by association, increased pain tolerance. Therefore, the purpose of this study was to explore (using binary classification), whether the implicitly measured demand can be determined using machine learning.

The results presented in this chapter indicate that, when using fNIRS data, a classification level of 86.4% was achieved when determining between easy and hard difficulty levels. As discussed, this classification was based largely on metrics relating to connectivity across fNIRS channels. Classification using fNIRS data was 31% more accurate in comparison to HR data (55.4%), which indicates that, although fNIRS data collection is a more expensive and invasive procedure, there are obvious benefits to using this data over alternate physiological data in regard to classification accuracy. Although the classification results observed in this study were promising for future development, we did not find statistical significance for the CPT. However, there was an observable trend, which indicated that pain tolerance was maximal in the medium and hard levels, compared to the easy or impossible levels.

Subjective self-report measures were only able to significant distinguish between the easy level, and all other levels.

Overall, the results of the study indicate that the use of fNIRS data yields a far superior classification than heart rate data. This study provided a successful method of pre-processing and classifying neurophysiology (fNIRS) and psychophysiology (heart rate) data to detect binary levels of game demand. The results have important implications for the future development of a neuroadaptive game. By focusing on the signals and measures that returned the highest classification accuracy in the design of a neuroadaptive game, it is possible to reduce processing time required for classification during a real-time protocol. These results are an improvement on previous studies where fNIRS data is specifically being used to determine game difficulty levels. A 2009 study found that classification between different levels of game demand was only 61.1% when using fNIRS data (Girouard et al., 2009). The results attained in this study are a promising improvement on this earlier work. As such, this improvement in classification rate may be due to the use of connectivity measures.

There were a number of limitations inherent in this study. Firstly, we had expected to find statistical significance between pain tolerance at every level of game difficulty. However, significance was only found between the Easy condition and all other conditions. One possibility for why this may have occurred is due to individual differences that are inherent with the CPT. Using the CPT allows no way to modulate the painful experience for the individual, which could have prevented a true representation of pain tolerance from being recorded. Three participants were removed from the CPT dataset prior to analysis due to short CPT immersion times, which only provided a sample size of 17 participants. This small sample size could have also negatively affected the significance of the CPT. As such, a larger sample size (and potentially an alternate pain protocol) could have improved the significance of the CPT results. Another issue that could have affected the CPT results is the overall effectiveness of the game distraction itself. It is possible that the game was not providing an adequate distraction. Although pilot testing was performed to inform the creation of the four levels of game difficulty, it is possible that there was still an inadequate range of demand available. It could have been the case that the four levels of demand were either too similar or too different to provide an immersive experience for the player at any point. The mechanics of the game could have also prevented an immersive experience if players found that the game as a whole was too complex to play, even at an easier level. It is also a possibility that the genre of the game was not immersive enough for some participants. As such, individual preference could have prevented the game from providing the desired experience. The game also did not allow for individual tailoring of the levels for each participant. For example, although we determined that we had created an ‘Easy’ game, we do not know that the player had the level of skill or experience necessary to consider this game easy. Equally, the ‘Impossible’ game may have still been too easy for some players. The overall inability to tailor each level of difficulty to the individual player meant that we had no way to ensure that each participant received the experience that we had aimed to provide.

Although the classification accuracies that were achieved in this study were promising, it is possible that the fNIRS montage that was chosen could have been improved. Feature selection indicated that connectivity measures provided the most relevant features for classification, which indicates that the most relevant fNIRS data could relate to connectivity between individual areas in the brain. For this reason, it should be considered that more accurate readings could be gathered if fNIRS data were monitored from individual areas of the brain that may work in association with each other. We should particularly consider choosing areas of the brain that are related to the experience of immersion, game demand, or attentional control, as discussed in Chapter 1.

The results gathered in this study have indicated that there is still further work that must be carried out before we can begin to develop a neuroadaptive game. It is posited that the most important step that should be taken next is to determine two or more individual areas of the brain that are affected by attentional control. As such, locating multiple cortical networks will enable us to take advantage of the effectiveness of connectivity analysis and that targeting areas that are directly related to attentional control will enable us to use connectivity analysis to determine how the demand of the game is affecting these different areas. As such, an adapted fNIRS montage may provide more accurate classification rates, which would improve the effectiveness of a neuroadaptive game. We will then focus on creating the neuroadaptive game, with a focus of researching previous work, particularly relating to time windows for data collection, and frequency of adaptation. We will also consider adaptations to the style of game used, in an effort to understand if this could improve the CPT tolerance rates and provide a clearer differentiation between levels of game difficulty.

The main findings of this study indicated that fNIRS data provided the most accurate classification rates, but that there was still further research required to determine the optimal fNIRS montage, which should be used for this work. This study also indicated that, in this case, the fluctuating difficulty of the game had not produced a statistically significant effect on pain tolerance, which we are interested in understanding further.

Chapter 4 - Study 2: Creation of Training Data for the Neuroadaptive Game

4.1 Introduction

Chapter 4 will focus on the collection and creation of training data which will later be used to inform the real time neuroadaptive system. The materials and methods used during data collection, and the processes carried out to attain the desired features, will be discussed throughout this chapter. Although the data collected in this chapter is primarily for use in Study 3, the classification of this Study 2 data will still be discussed, with a view to using the data in real time.

The purpose of this study was to collect a dataset that could be used to train a real-time classifier, which would be used to create a neuroadaptive game. We wished to run this experiment with an fNIRS montage that reflected work on attentional pathways in the cortex. We wished specifically to focus on the top-down attentional pathway, which relates to the gameplay task, and the bottom up pathway, which reflects the experience of pain (Eccleston & Crombez, 1999; Legrain et al., 2009; Torta, Legrain, Mouraux, & Valentini, 2017).

Attention is a limited resource, and a participant who is presented with more than one stimulus at any given time must make a choice on how to divide their limited attention pool. This decision, however, is not always a conscious one. Pain is an interruptive function, which means that a painful experience is capable of drawing attention away from task stimuli. However, varying characteristics of both the distraction task and the painful stimuli can affect how attention is shared between the two said tasks. The theory that attention must be shared between tasks implies that there is an interaction occurring between the two areas of the brain, which are involved in the different tasks. Attention is dynamically switched between top-down and bottom-up sources, depending on which process is more salient (Johnson, 2005).

When a person experiences pain, their brain will usually reorient to this interruptive stimulus. Reorienting refers to the implicit reaction that occurs within the brain when a person is faced with noxious stimuli. The purpose of reorienting is to preserve the self, that is, to avoid or remove oneself from situations that may cause, or are causing, pain. The two areas of the brain where we expect to see an interaction between the top-down and bottom-up orienting systems are the ventral frontoparietal network and the dorsal frontoparietal network. The dorsal frontoparietal network is made up of the Intraparietal Sulcus (IPS), the Superior Parietal Lobule (SPL) and the right and left Frontal Eye Fields (FEF). These areas work in conjunction to form the top-down control mechanism. The ventral frontoparietal network, which relates to bottom-up activity such as a painful experience, consists of the Temporoparietal Junction (TPJ) and the Ventral Frontal Cortex (VFC). When a person experiences a painful experience, the right hemisphere dominant ventral frontoparietal network can interrupt activity, which is not related to the painful experience, in order to allow the focus of a person to switch from task related activity to attending to the pain that they are experiencing. The dorsal frontoparietal network enables a person to make the connection between stimuli and response. For example, if a person puts their foot into a tank of cold water, which causes pain, then the dorsal frontoparietal network is responsible for making the link between the presence of the foot in the water, and the experience of pain. The dorsal pathway causes re-orientation to any salient interruptive stimulus (noise from phone, person talking), not just painful stimuli. The salience of interruptive stimuli is partially related to the relevance of the specified stimuli, and partially related to the current level of task demand. This means to say that, stimuli that may not be considered interruptive whilst a person is performing a demanding task could be considered interruptive if the person is performing a non-demanding task, or, if said person is not entirely focused on the task (for example, due to boredom or overload relating to task demand). When the brain is at rest, the individual networks (Dorsal and Ventral) perform independently to each other. However, when a person is focused, the ventral network (which interrupts the current task and begins the reorienting process) is subdued to prevent distractions from effecting the main task goal. This research indicates that brain activity from the ventral frontoparietal network could monitored for the purpose of determining engagement specifically during the experience of pain. Lessened activity in the ventral frontoparietal network could indicate a high level of focus, whereas increased activity may indicate a lack of focus, or that the painful experience has reached the point where a person finds it unbearable (Arrington, Carr, Mayer, & Rao, 2000; Bundesen, 1990; Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Corbetta, Patel, & Shulman, 2008; Kincade, Abrams, Astafiev, Shulman, & Corbetta, 2005).

In recent years, further research has indicated that brain functions do not occur in individual areas of the brain on a task-by-task basis, but rather that networks of brain areas are responsible for producing brain function (Fox et al., 2005; Laird et al., 2013). Exploring relationships between brain areas within a brain network could provide results that examine both the individual activities of each network, and also any interactions that may occur between multiple networks. For this reason, it was decided that the fNIRS montage for study 2 should focus on both the ventral attention network and the dorsal attention network (Corbetta et al., 2008; Torta et al., 2017). The dorsal and ventral systems are almost completely segregated, (although they do share the common area of the TPJ) but there is evidence to suggest that, although the two dorsal networks do not interact, they may be linked through the prefrontal cortex and share similar spontaneous activity (Fox, Corbetta, Snyder, Vincent, & Raichle, 2006). Areas within the ventral and dorsal attentional network include the Intraparietal Sulcus, Superior Parietal Lobule, Temporoparietal Junction, Frontal Eye Fields, Inferior Frontal Gyrus and Medial Frontal Gyrus. These areas can be mapped onto the 10/20 EEG system, and then used to determine a suitable fNIRS montage that could effectively collect data from these areas.

The aim of this study is to manipulate task demand alongside the presence of experimental pain (similar to Study 1). Yet, we hypothesise that in Study 2 we will i) observe greater dorsal activation in response to increased levels of game demand and ii) be able to categorise game demand using fNIRS connectivity measures, which will provide a higher level of classification accuracy compared to using measure of game score alone.

4.2 Methodology

4.2.1 Design

Study 2 was a two-factor repeated measures study, consisting of three levels of game demand (Easy, Hard and Impossible) that are experienced with and without experimental pain, administered via the cold pressor test. The six games were delivered in randomised order.

4.2.2 Participants

The study consisted of 23 (11 female) participants. Participants were aged between 18 and 34 *(M = 24, SD = 2.6).* Four participants were excluded from this dataset due to consistently short immersion times during the cold pressor test (< 15 seconds). It was determined that these participants did not have a sufficient baseline pain tolerance to experience an increase in pain tolerance during gameplay. Prior to the beginning of the experiment, each participant gave informed consent and completed a demographic questionnaire to ensure that they were suitable to continue with the experiment. The exclusion criteria for participants were identical to the experiment reported in the previous chapter. The procedure for the experiment and data collection protocol was approved by the Liverpool John Moores University (LJMU) Research Ethics Committee and the experiment was conducted in accordance with the recommendations of the LJMU Research Ethics Committee.

4.2.3 Materials

4.2.3.1 Ribbon Rush

Ribbon Rush is a collision avoidance driving game (as described in Chapter 2). The goal of the game is to score as high as possible. A high score can be gained by avoiding collisions with non-playable character (NPC) enemy cars, which are driving in the opposite direction to the player car. If the player collides with an enemy car, then points are deducted from their score. A high score is more quickly achieved if the speed of the game is increased, but an increase in game speed also makes for a more difficult gameplay experience. The nature of the game Ribbon Rush meant that a score was generated every 2 seconds. This score was cumulative and was increased automatically throughout gameplay. The score was reduced if the player car was involved in a collision with an NPC car. Features created from the game score data were Mean, Peak and Variance.

4.2.3.2 Experimental Pain

The Cold Pressor Test (CPT) was used to gather a behavioural measure of pain tolerance for each of the three games, as well as a baseline measurement of pain tolerance.

4.2.3.3 Subjective Measures

Subjective immersion was measured via the NASA Task Load Index (TLX) and the Motivation scale.

4.2.3.4 Functional Near Infrared Spectroscopy

fNIRS data was collected using an Oxymon Mk III, which recorded changes in optical density from the participant's brain. This was used as a measure of neuroactivity. The participant wore a cap consisting of 6 channels (6 transmitters and 6 receivers) on their head. The areas where channels were created are outlined below in Figure 30, according to the 10/20 system:

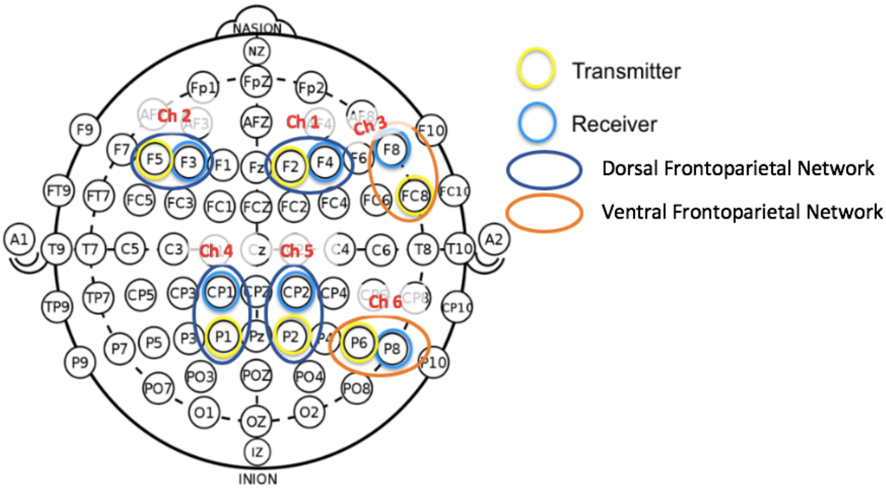


Figure 30 - fNIRS montage

Although the Oxysoft software provides a selection of fNIRS montages that are available for use in data collection, we were not satisfied that any of these montages were fit for our purpose. For this reason, we designed our own montage with the intention of targeting the dorsal and ventral attentional systems simultaneously. Figure *30* above indicates the channels of the chosen fNIRS montage, which are related to these two individual networks. As indicated by Figure 30 and Table 5*,* channels 1, 2, 4 and 5 related to Top-Down control, whereas channels 3 and 6 relate to stimulus driven control.

We chose to make adjustments to the fNIRS montage for Study 2. Although we had achieved results which we considered to be good enough for real-time adaptation in Study 1, further research indicated that connectivity measures may provide a more accurate classification rate in a real time setting. Although connectivity measures could have been created from our Study 1 montage, the adapted montage which we created for Study 2 was more focused on connectivity between brain areas. This montage allowed us to gather data independently from the dorsal attentional systems and the ventral attentional systems – and also create features which provided insight into connections made between these areas. Although the dorsal and ventral systems relate to different brain activity (top down activity such as directed attention and bottom up activity such as pain response), there is correlation between these areas when a task is switched (for example, when switching between focusing on a video game to attending to pain). We hypothesised that targeting these independent areas and the correlations between them would provide a more thorough insight into the activity occurring in the brain throughout the study.

Table 5 - Channels and their relation to the DFN/VFN

Dorsal Frontoparietal Network	Ventral Frontoparietal Network
Channel 1 – Right Frontal Eye Field	Channel 3 – Right VFC
Channel 2 - Left Frontal Eye Field	Channel 6 – Right TPJ
Channel 4 – Left Ips	
Channel 5 – Right Ips	

4.2.4 Procedure

Before the experiment began, each participant completed a familiarization with both the game, and the CPT. This was to ensure that the data that was collected was an accurate representation of the participants pain tolerance and gaming ability and was not affected by being unfamiliar with the stimulus. Following familiarization, the participants completed a timed CPT without a distraction to measure their pain tolerance. This control was carried out to ensure that an accurate measurement of the differences in pain tolerance between the two game conditions could be established. This CPT baseline was gathered prior to each game and pain condition, for a total of three timed CPT baselines. Participants played each level of game difficulty twice, once whilst experiencing the CPT, and once without. This provided a total of six game conditions gathered for each participant. Participants were always instructed to alternate the foot that they used in the CP. For example, if the familiarization was carried out with the left foot, then the pain tolerance baseline was recorded using the right foot, and then the game and pain condition was carried out using the left foot again. This was to ensure that the foot did not either become oversensitive to pain (and therefore reduce overall pain tolerance) or become numb due to the cold water, and therefore provide a false representation of increased pain tolerance.

The experiment began with a 30 second baseline, where the participant was asked to remain calm and prepare them self for the experiment. After this baseline period, a baseline CPT measurement was recorded. This was to ensure that familiarisation to the CPT did not affect the overall pain tolerance results. Following this baseline, participants began to play the first level of difficulty, which had been chosen for them, whilst experiencing the CPT. The game difficulty order was chosen at random and was kept blind from the participants until the end of the experiment session. Each game session lasted for three minutes. When the participant was experiencing the game and pain condition, they were instructed to place their foot in the water as the experimenter started the game. Participants were instructed to keep their foot in the CPT until they felt that they had to remove it. The participants continued to play the game for the full three minutes, regardless of when they removed their foot from the water. Following the completion of the game and pain condition, the participant was given the opportunity to dry their foot and warm it up, to ensure that any pain they may still be experiencing from the CPT was alleviated. The participant was then asked to complete the TLX questionnaire to establish their level of engagement with the game that they had just played. Following the completion of questionnaires, the participant played the same level of game difficulty again, but this time without experience pain induced by the CPT. This condition also lasted for three minutes. Following this, the participant was instructed to complete the TLX for a second time, to determine whether the inclusion of pain had had an effect on their experience of pain.

This condition procedure was repeated twice, for a total of three conditions per experimental session. This allowed for the participant to experience each level of game difficulty. Once the participant had completed all conditions, they were debriefed, and the experiment concluded. The results of the study, including CPT times and level adaptations which were made during each game, will be discussed further in the following section.

4.2.5 Signal Processing Pipeline

As Study 2 was carried out in preparation for creating a real-time neuroadaptive game, it was decided that this study would be carried out without the use of the AMARA algorithm. This would enable us to determine whether the fNIRS data could still be filtered efficiently without the use of accelerometer data. Disregarding the AMARA filter would speed up the pre-processing process during real-time adaptation. Figure 31 below depicts a flow diagram of the pre-processing steps that were applied to the fNIRS data.

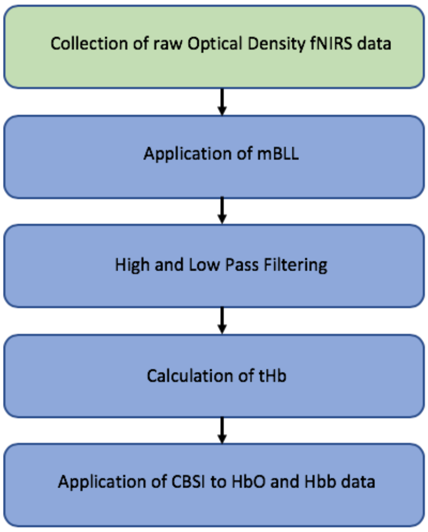


Figure 31 - Flow diagram of filtering and pre-processing techniques used to prepare fNIRS data for feature creation

The first step in the fNIRS data pre-processing was to apply the Modified Beer Lambert Law (mBLL) to the raw optical density (OD) signals, in order to create Oxygenated and Deoxygenated haemoglobin data. Manually applying the mBLL allows us to determine our own baseline period, which will produce a more accurate representation of haemoglobin change based on exposure to stimuli. Following application of the mBLL, data were high and low pass filtered to remove underlying noise, which may be present in the signal. High and low pass filtering was carried out using a 6th order Chebyshev filter with 0.1 Hz and 0.3 Hz cut-off frequencies respectively. Filter parameters where adjusted slightly between Studies 1 and 2 due to the removal of the AMARA filtering technique. AMARA was used in Study 1 but was not used in following studies due to the lack of real-time filtering capability present in the AMARA filter. Following basic filtering, total haemoglobin (tHb) was calculation through addition of the HbO and HHb signals. tHb was calculated at this stage to prevent the Correlation Based Signal Improvement (CBSI) filter from distorting the tHb signal. CBSI was applied following tHb calculation to remove any remaining artefacts (particularly from head movement or slippage), which may have been present in the fNIRS data. At this stage, fNIRS data were pre-processed and feature extraction could begin.

Features were extracted from both the Game Score data and the fNIRS data in 10 second windows. Data were windowed to reduce the overall quantity of data to produce a faster classification. The purpose of reducing the dataset and increasing classification speed was to ensure that data could be classified both quickly and accurately. If an acceptable level of classification accuracy could be attained from this data, we could use the outlined processing pipeline in real time for further studies.

4.3 Results

The results of this study were analysed with reference to measures of pain and subjective mental workload, as well as game difficulty classification. The results of the Pain section discuss the participants objective responses to pain, which is measured by the amount of time that they are able to keep their foot within the CP tank. Subjective mental workload was measured via the TLX and Motivation scales, which were scored post-hoc and analysed. Game difficulty classification relates to the ability of both fNIRS data and Game Score data (independently and combined) to classify the level of game difficulty that the participant was playing. Subjective mental workload was measured to determine whether there was a significant difference between the three levels of game difficulty, and to determine whether participants subjective immersion had an effect on their pain tolerance.

4.3.1 Pain

The three CPT baseline times, which were recorded prior to each game condition (easy, hard and impossible), were averaged to create a mean baseline pain tolerance time. The mean pain tolerance score was then compared to the pain tolerance score, which was recorded during each game condition, to see if difference in pain tolerance were represented within the data. Figure 32 below is a representation of these pain scores averaged across all participants.

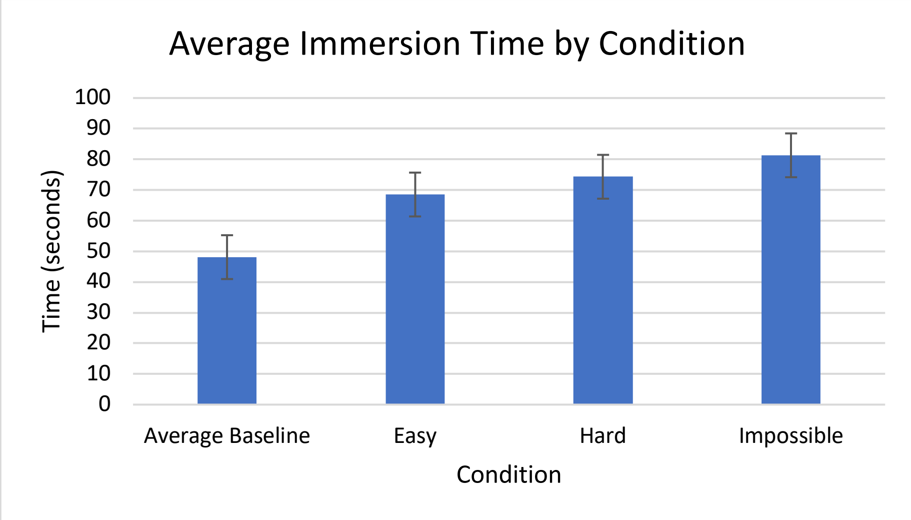


Figure 32 - Average CPT results

A one-way ANOVA was performed on the CPT results to determine the effect of game difficulty on pain tolerance. The analysis revealed a significant main effect for demand [F (3,16) = 4.89, p<.01, η^2=0.48]. Post-hoc tests revealed that there was a significant difference between the baseline condition and all other conditions (p = .03).

4.3.2 Subjective Mental Workload & Motivation

Table 6 below displays the average results of the TLX and Motivation questionnaire for the 19 participants whose data was included in this study. These results were gathered after the Game and Pain condition, and therefore reflect the participants perception of game difficulty level whilst also experiencing pain.

Table 6 - Average TLX and Motivation questionnaire results

	Easy		Hard		Impossible	
	TLX	**Motivation**	**TLX**	**Motivation**	**TLX**	**Motivation**
Mean	3.45	37.68	4.59	40.21	5.51	41
Std	1.35	4.89	1.62	4.94	1.7	4.37

Table 6 indicates that both participants mental workload and motivation increased in line with the game difficulty. On average, participants did not experience a drop-off of either motivation or mental workload, even during an Impossible game condition. A one-way ANOVA was performed on the TLX results to determine the effect of game difficulty on mental workload. The analysis revealed a significant main effect for demand [F (2,17) = 22.28, p<.00, η^2 = 0.72]. Post-hoc tests revealed that there was a significant difference between all three game conditions (p < .01), i.e. subjective workload was higher for impossible demand compared to the other two conditions and hard demand evoked higher workload compared to easy demand.
A one-way ANOVA was also performed on the Motivation results to determine the effect of game difficulty on subjective motivation. The analysis revealed a significant main effect for demand [F (2,17) = 4.18, p<.03, η^2 = 0.33]. Post-hoc tests revealed that subjective motivation was significantly lower for Easy compared to the higher demand conditions (p < 0.5).

4.3.3 Analysis of Effect of Game Demand on Individual and Combined Channels

Connectivity measures were chosen for classification as it had been previously established (see Chapter 3) that these were the most relevant. Although we understand that these measures are more relevant, it is important to explore why this may be occurring. The montage that was chosen for this study was specifically selected due to the expected communications between the dorsal and ventral networks. To explore whether activity was occurring between these networks (specifically in relation to game demand), the data from these channels were analysed to determine whether there were changes in activation during each of the levels of gameplay. A series of one-way ANOVAs were performed on this data to determine whether game difficulty had a significant effect on neurovascular response from any individual channel. Tables 7 – 8 below indicate the results of this analysis, and Figure 33 represents the significant effects as a bar graph.

Table 7 - Pairwise ANOVA results for HbO data

Channel	F(2,36)	Sig
1 (R FEF)	0.41	0.67
2 (L FEF)	1.5	0.25
3 (R VFC)	0.58	0.57
4 (L IPs)	0.70	0.51
5 (R IPs)	0.43	0.66
6 (R TPJ)	0.45	0.65

Table 8 - Pairwise ANOVA results of tHb data

Channel	F(2,36)	Sig
1 (R FEF)	0.03	0.97
2 (L FEF)	0.21	0.81
3 (R VFC)	0.90	0.43
4 (L IPs)	**4.12**	**0.04**
5 (R IPs)	0.45	0.65
6 (R TPJ)	0.33	0.72

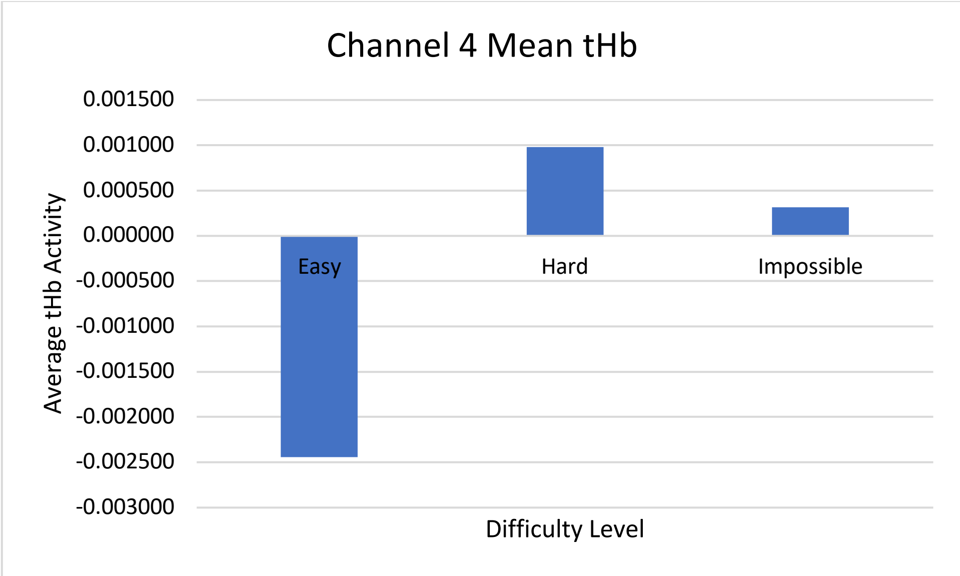


Figure 33 - Bar chart illustrating mean tHb for Channel 4 at each level of difficulty

The analysis revealed that the only channel were game difficulty had a significant effect was Channel 4 – and this significance only occurred in the tHb. This significance is indicated in Table 8. Pairwise comparisons of all other channels and feature types (HbO or tHb) found no statistically significant difference in activity depending on difficulty level ($p>.05$). Statistical significance for tHb in Channel 4 was only found between easy and impossible levels ($p = .03$) (Figure *33)*. Differences in tHb at Channel 4 were not significant between easy and hard ($p = .3$) or hard and impossible ($p = 1$).

4.3.4 Feature Selection for Classification

Feature selection was carried out in order to reduce the feature set, which can lead to improved classification accuracy and classification speed. Before feature selection can be carried out, all features must be collated. Features included in feature selection contain descriptive statistics and correlation results. Following the collation of features, this data was subjected to the RELEIFF algorithm, which determines the relevance (weight) of features, and identifies the point in which features are no longer capable of influencing classification. The features that were selected for classification can be seen below in Table 9.

Table 9 - Selected features as defined by the RELIEFF algorithm

HbO	Weight	tHb	Weight
Ch1x4	0.0124	Ch1x4	0.0124
Ch1x5	0.0142	Ch1x5	0.0142
Ch1x6	0.0052	Ch2x4	0.0109
Ch2x4	0.0110	Ch2x5	0.0143
Ch2x5	0.0143	Ch3x4	0.0142
Ch2x6	0.0124	Ch3x5	0.0180

Ch3x4	0.0142	Ch5x6	0.0134
Ch3x5	0.0180		
Ch3x6	0.0002		
Ch4x6	0.0075		
Ch5x6	0.0134		

Figures 34-36 below illustrates correlation, using Pearson's Correlation Coefficient, in relation to the changes in HbO coactivation between Channel 1 (Right Frontal Eye Field (R FEF)) and Channels 4 (Left Intraparietal Sulcus (L IP)), 5 (Right Intraparietal Sulcus (R IP)) and 6 (Right Temporoparietal Junction (R TPJ)). The changes in HbO coactivation between Channel 2 (Left Frontal Eye Field (L FEF)) and Channels 4 (L IP), 5 (R IP) and 6 (R TPJ) are also depicted.

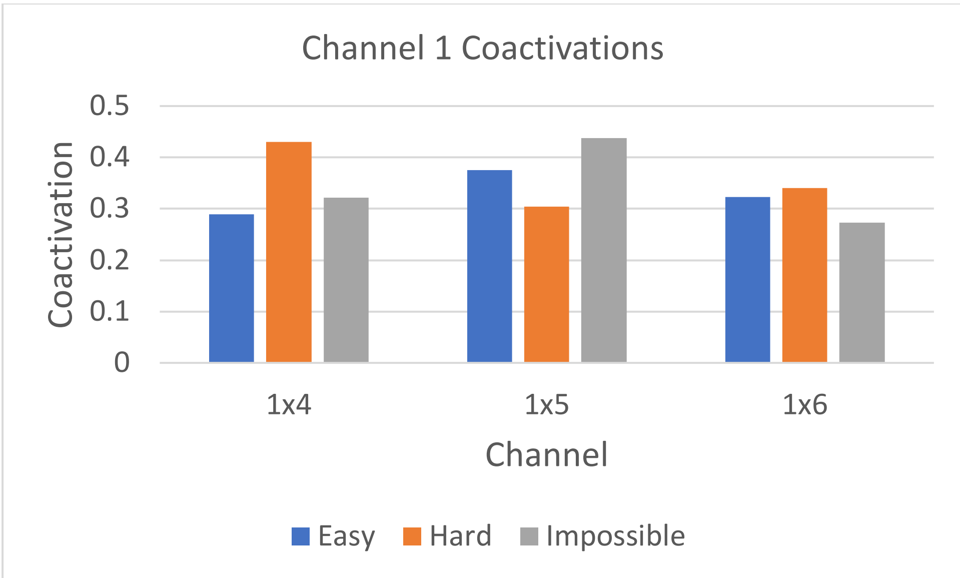


Figure 34 - HbO Coactivations for Channel 1x4, 1x5 and 1x6

For correlations involving Channel 1 (L FEF) (Figure 34), 1x4 increased for Hard demand compared to the other two conditions, whereas this pattern was reversed for 1x5. 1x6 revealed a reduction of coactivation when game demand was impossible. On the right hemisphere (1x5), increased connectivity can be observed during impossible demand, whereas connection between right FEF and Left IPs (1x4) peaked during hard demand. With respect to dorsal-ventral connectivity between right FEF and TPJ (1x6), we see a decline during impossible demand compared to the other two conditions.

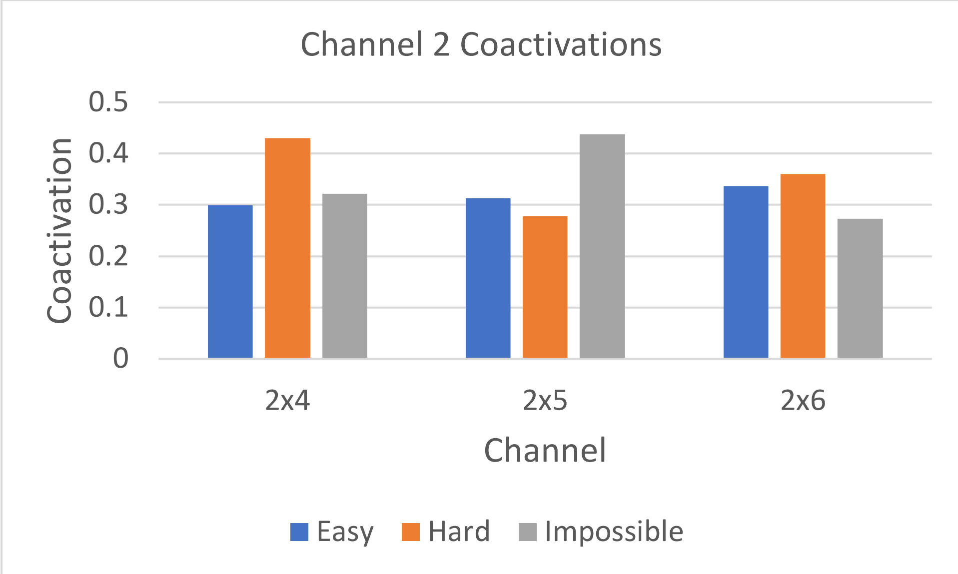


Figure 35 - HbO coactivations for Channel 2x4, 2x5 and 2x6

For correlations involving Channel 2 (R FEF) (Figure 35), 2x4 increased for Hard demand compared to the other two conditions, whereas this pattern was reversed for 2x5. 2x6 revealed a reduction of coactivation when game demand was impossible. On the left hemisphere (2x4), increased connectivity can be observed during hard demand, whereas connection between Left FEF and Right IPs (2x5) peaked during impossible demand. Dorsal-ventral connectivity between left FEF and TPJ (2x6) increases between easy and hard demand but decrease again for impossible demand. Activity at 2x6 is at its lowest point during impossible demand.

Figure *36* below indications patterns of coactivation between Channel 3 (Right Ventral Frontal Cortex (R VFC)) and Channels 5 (R IP) and 6 (R TPJ). For correlations involving Channel 3 (R VFC), 3x5 increased for Impossible demand compared to the other two conditions, whereas there was a gradual reduction in activation at 3x5 as the difficulty of the game increased. Dorsal-ventral activity at the Right VFC and Right IPs (3x5) increases gradually as the difficulty of the game increases and is at its highest point during the impossible condition. Connectivity between the Right VFC and Right TPJ (3x6) in the ventral network indicates an opposite pattern of connectivity, where activation is highest during easy demand, and decreases between both easy and hard, and hard and impossible.

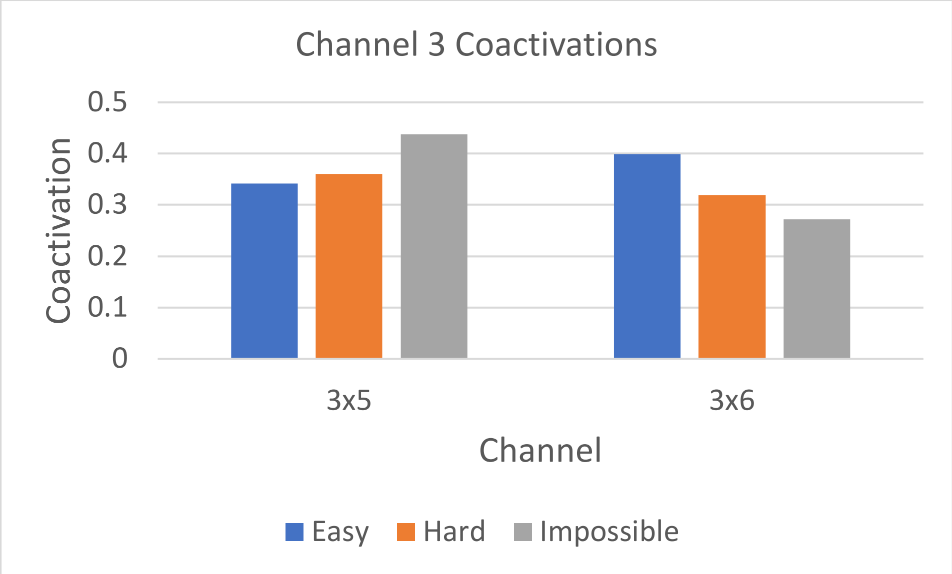


Figure 36 - HbO coativations for Channel 3x5 and 3x6

Figure 37-38 below indicate patterns of coactivation between Channel 4 and Channel 6, and Channel 5 and Channel 6.

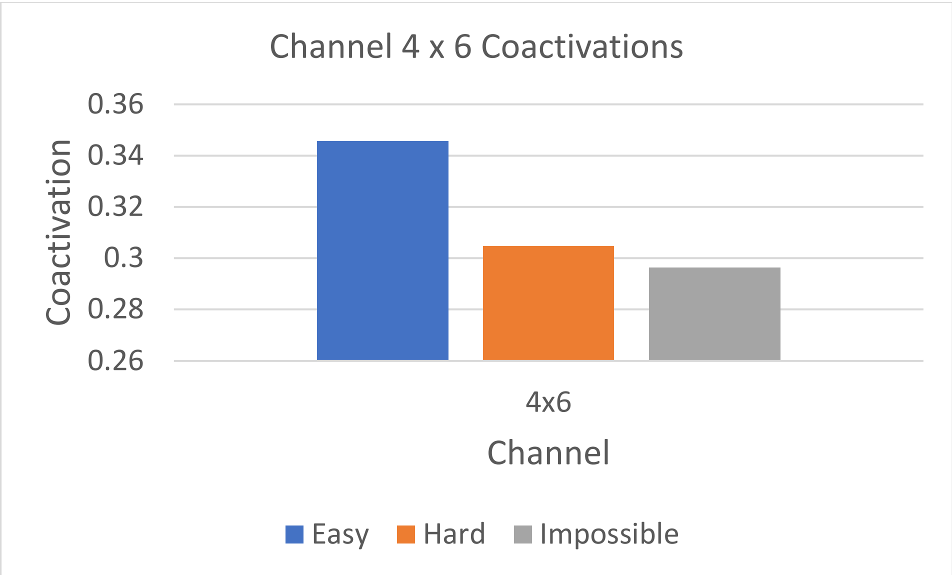


Figure 37 - HbO coactivations for Channel 4x6

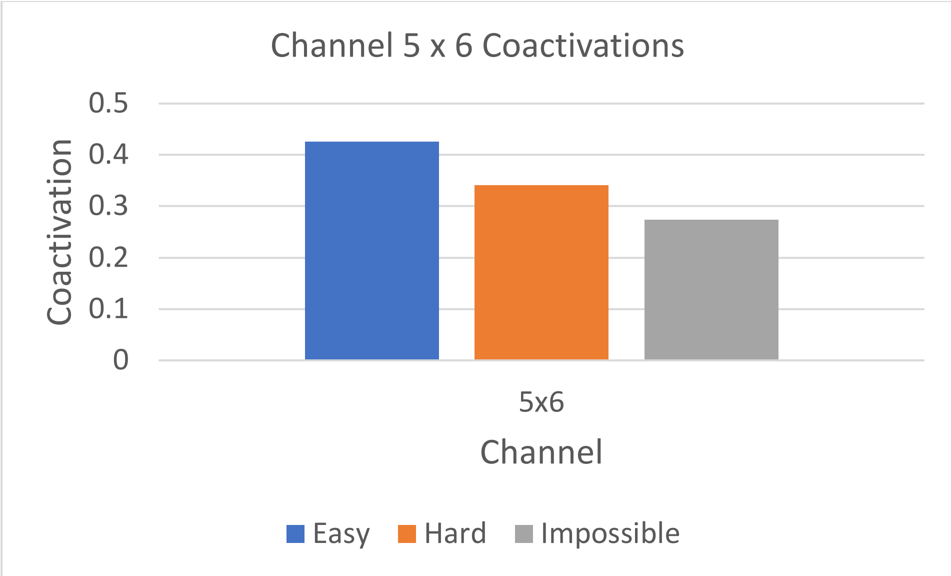


Figure 38 - HbO coactivations for Channel 5x6

For correlations involving Channel 4 (L IPs), a linear reduction in coactivation in 4x6 (Figure 37*)* was found as the difficulty of the game increased. Similarly, a linear decrease in coactivation in 5x6 was found as the difficulty of the game increased. For correlations involving Channel 4 (L IPs), a linear reduction in coactivation in 4x6 (Figure 38) was found as the difficulty of the game increased.

4.3.5 Classification

Classification was carried out in a binary format to include all three levels of difficulty. To compare each of the difficulty levels entirely, binary classification was carried out three times: Easy v Hard, Hard v Impossible and Easy v Impossible. The RELIEFF algorithm was used to determine relevant features for classification from this feature set. Feature selection determined that 18 fNIRS features would be most relevant for classification. These features can be seen above in Table 9*:*

The data used for classification purposes consisted of the Game data; therefore, each participant had an equal length dataset. For this reason, data balancing was not required before classification could be carried out. A current game score was collected every 2 seconds during this experiment. This game score was classified alone to determine if the score of the game could be used to determine game difficultly level. These classifications can be seen below in Table 10.

Table 10 - Classification results from Game Score data

	Easy v Hard		**Easy v Impossible**		**Hard v Impossible**	
Measure	**LDA**	**SVM**	**LDA**	**SVM**	**LDA**	**SVM**
Acc	**65.36%**	63.03%	69.16%	**88.59%**	73.26%	**82.47%**
BER	**34.21%**	36.57%	30.75%	**11.35%**	27.03%	**17.47%**
F1	55.50%	**62.78%**	58.57%	**88.04%**	64.22%	**81.47%**

It can be determined from the results illustrated in Table 10 that classification between Impossible levels of difficulty and the two alternate levels (Easy and Hard) was acceptable (Acc = 88.59% and 82.47% respectively for Easy v Impossible and Hard v Impossible). However, classification accuracy between Easy and Hard levels was poor (65.36%). SVM was superior to LDA for both Easy v Impossible (SVM Acc = 88.59, LDA Acc = 69.16%), and Hard v Impossible (SVM Acc = 82.47%, LDA Acc = 73.26%), but LDA was superior for Easy v Hard. The most accurate classification accuracy was found between the Easy v Impossible condition, where a higher accuracy and F1 and lower BER were found. This indicates that, as expected, classification is more accurate between the most significantly different levels of game demand. The observance of reduced classification capability in the lower range of demand indicates that, where game score classification to be used in real time, we should expect that inaccurate adjustments would be made in this range due to the lack of sensitivity in the game score classification.

fNIRS data were also classified in 10 second windows to determine if a stronger classification accuracy could be found in the fNIRS data as opposed to the game score data. The results of this classification can be seen below in Table 11.

Table 11 - Classification results from fNIRS data

	Easy v Hard		**Easy v Impossible**		**Hard v Impossible**	
Measure	**LDA**	**SVM**	**LDA**	**SVM**	**LDA**	**SVM**
Acc	81.00%	**86.40%**	82.60%	**88.74%**	77.34%	**91.09%**
BER	18.81%	**13.62%**	17.07%	**11.22%**	22.45%	**8.78%**
F1	80.19%	**86.42%**	82.12%	**88.60%**	77.01%	**91.07%**

In relation with the classification of game score, the poorest classification accuracy is found between the Easy and Hard levels of difficulty when fNIRS data is used for classification (LDA Acc (81%, SVM Acc 86.4%). However, fNIRS data is still providing a vastly improved classification level even during the Easy v Hard condition (86.4% for fNIRS data compared to 63.03% for Game Score data). Classification accuracy and F1 scores are all over 85% when fNIRS data is used for classification, however, BER is still above 10% in both the Easy v Hard and Easy v Impossible conditions.

Table 11 indicates that fNIRS data can yield a more accurate classification than Game Score data. However, we were also interested to determine whether a combination of the two data types could provide any improvement on these classification rates. The results of the combined fNIRS and Game Score classification can be seen in Table 12 below. The window used for classification when fNIRS and Game score data were combined remained at 10 seconds. This meant that 100 instances of fNIRS data and 5 instances of game score data were windowed to provide one value for each dataset.

Table 12 - Classification measures from combined fNIRS and Game Score data

	Easy v Hard		Easy v Impossible		Hard v Impossible	
Measures	**LDA**	**SVM**	**LDA**	**SVM**	**LDA**	**SVM**
Acc	83.04%	**85.38%**	85.95%	**94.88%**	84.65%	**94.88%**
BER	16.85%	**14.87%**	13.99%	**5.30%**	15.36%	**5.09%**
F1	82.68%	**85.20%**	84.99%	**94.70%**	83.58%	**94.84%**

There is a marked improvement in classification accuracy when a combination of game score and fNIRS data is used for difficulty level classification. The improvement is most obviously observed in the Easy v Impossible (94.88% for Combined data compared to 88.74% and 88.59% for fNIRS and Game Score data respectively) and Hard v Impossible classifications (94.88% for Combined data compared to 91.09% and 82..47% for fNIRS and Game Score data respectively), where accuracies are approaching 95% and the BER is 5%. Although these results are promising, the classification between Easy v Hard is still somewhat problematic (85.38%), especially considering the 14% BER. These consistently poorer results, which can be observed in an Easy v Hard condition, indicated that there may be a higher level of variability in both fNRS data and game score data between these levels.

A 3x3x2 Repeated Measures ANOVA was carried out to determine the main effects of game difficulty, classification measures and classification method. We carried out this ANOVA to determine whether there was statistical significance between a) different measures (fNIRS, Game Score and fNIRS & Game Score combined), b) different difficulty levels (Easy v Hard, Easy v Impossible, Hard v Impossible), and c) different classification techniques (LDA and SVM). To carry out this ANOVA, the F1 scores gathered from classification were used.

The results of the 3x3x2 Repeated Measures ANOVA revealed that there was statistical significance in all three of these areas, as indicated below:

- A) Measures – [F(2,17) = 20.32, p = <.01, eta = .705]

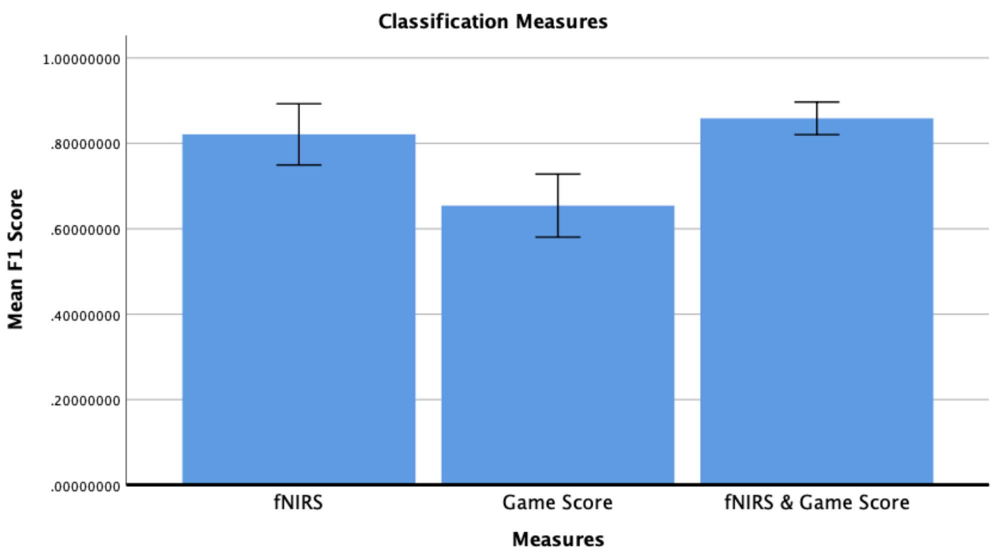


Figure 39 - Mean F1 Score for Classification Measures

According to pairwise comparisons, F1 score for Game Score was significantly lower than either fNIRS or fNIRS & Game Score (p<.01). Figure 39 provides a visualisation of these findings.

- B) Levels – [F(2,17) = 6.56, p = <.01, eta = .436]

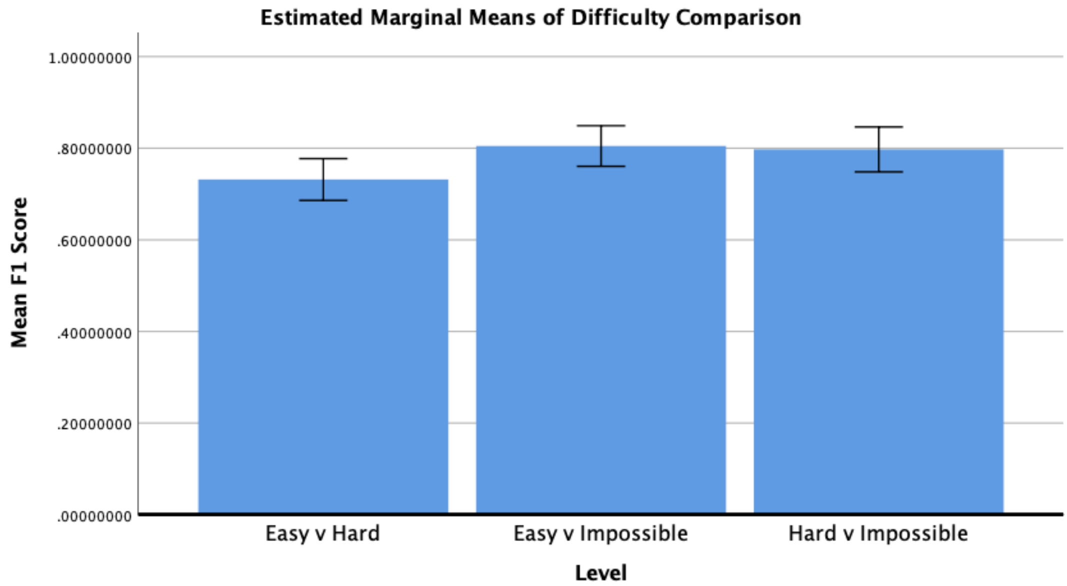


Figure 40 - Mean F1 Score for Difficulty Comparison

Figure 40 indicates that a larger difference in level comparison was found between Easy v Hard and both Easy v Impossible and Hard v Impossible. Pairwise comparisons of the F! score for level comparison revealed that F1 was significantly higher for Easy v Impossible and Hard v Impossible compared to Easy v Hard (p <.01).

- C) Classification – [F(1,18) = 38.59, p = <.01, eta = .682]

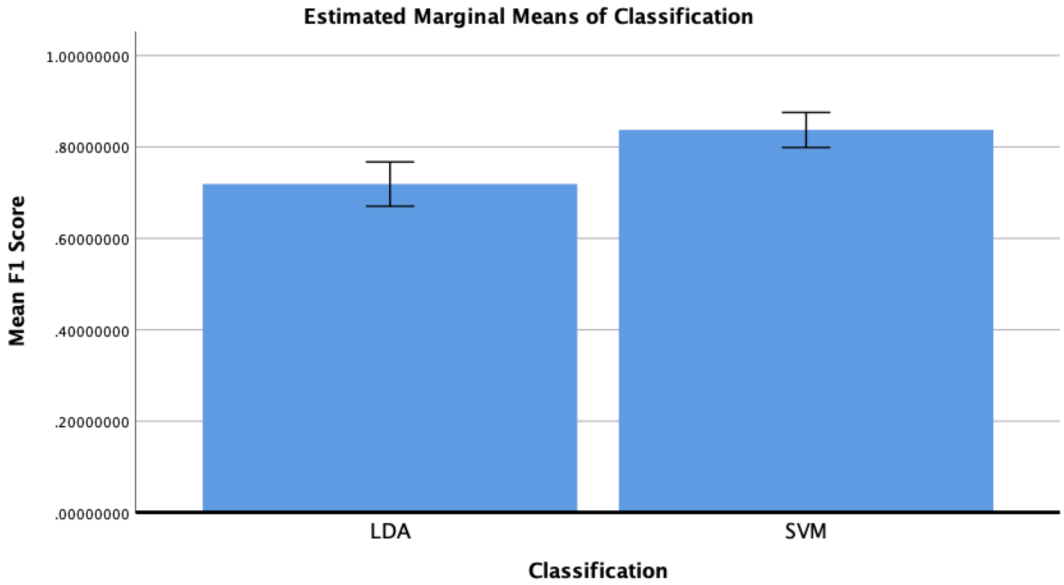


Figure 41 - Mean F1 Score for Classification

Figure 41 indicates that the SVM classification produced significantly higher F1 score compared to LDA. Pairwise comparison of the F! score indicates that this comparison was significant (p <.01).

Considering the significant interaction between the Measures and Classification, Figure 42 appears to indicate that that the SVM classification technique is superior to the LDA classification technique across all difficulty comparisons.

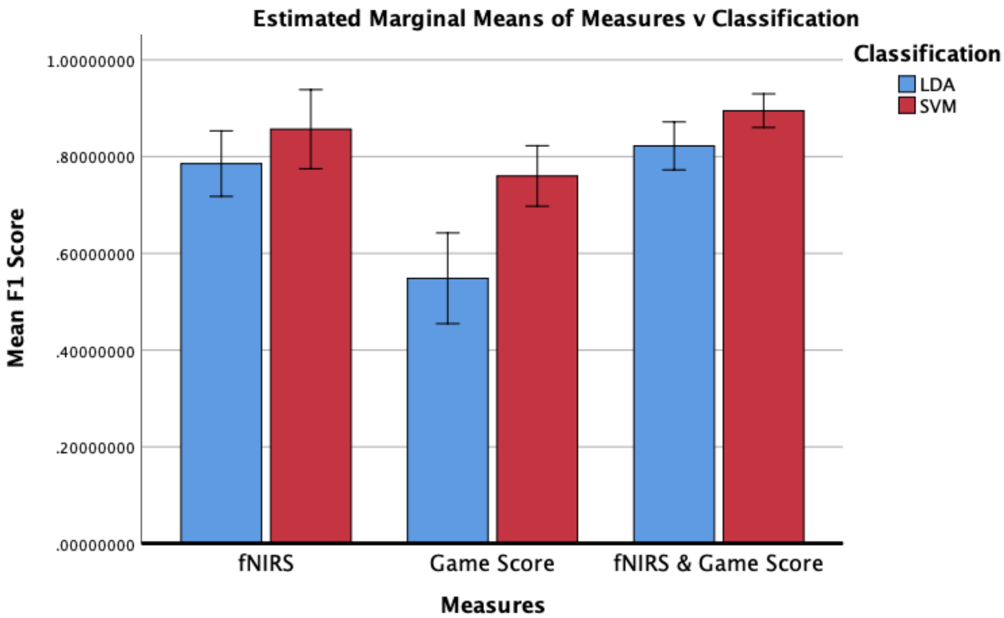


Figure 42 - Mean F1 Score for Measures v Classification

To explore this result further, a series of paired sample t-tests were carried out, which revealed that fNIRS data is significantly more accurately classified than game data when an LDA is used ($p = <.01$), but a combination of fNIRS & Game Score data is not significantly more accurate than fNIRS data alone ($p = <.125$). However, a combination of fNIRS data and Game Score data is significantly more accurately classified than Game Score data alone ($p = <.01$).

SVM produced consistent performance over all three groups of measures, whereas F1 score was significantly lower for LDA models using Game Score compared to both fNIRS and fNIRS & Game.

This result indicates that a combination of fNIRS & Game provides no further improvement against using fNIRS results alone for SVM classification. These results indicate that the significant interaction comes from the comparison of fNIRS & Game and game alone, where the F1 score for SVM is significantly higher ($p = <.01$).

To explore the interactions within the data further, we also carried out a series of paired sample t-tests to determine how the accuracy of each measure (fNIRS, Game Score, fNIRS & Game Score) was affected by the level comparison which was used (EvH, EvI, HvI).

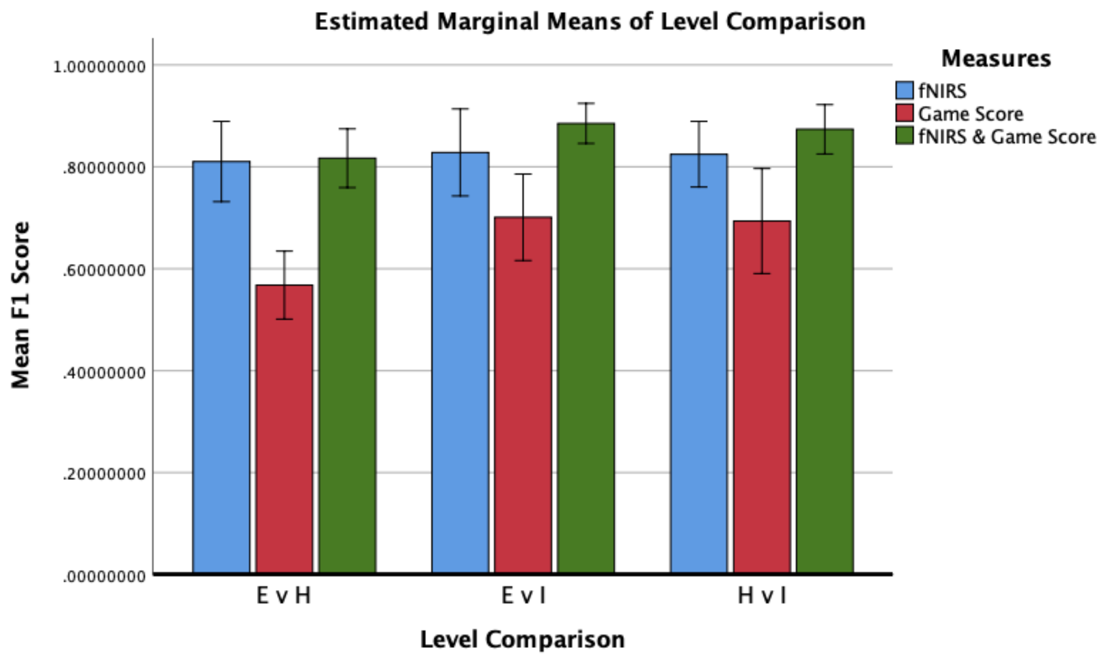


Figure 43 - Mean F1 Score for Level Comparison v Measures

The results of the t-tests carried out on this data revealed a series of interesting interactions. It was revealed that using fNIRS & Game Score data combined is always statistically more accurate than using Game Score data alone ($p = <.01$). This indicates that, regardless of the level comparison or classification technique, a combination of the two data types always provided a more accurate classification compared to simply using the Game score. This is to be expected, as the classification results already show that Game Score is the weakest classification measure overall. It was also revealed that fNIRS data and fNIRS & Game Score were never significantly different from each other regardless of level comparison, or classification type ($p = >.1$).

Further analysis of the interaction between Level Comparisons and Classification Measures revealed that the F1 score for Game Score was only significantly worse than either fNIRS or fNIRS & Game Score during the Easy v Hard level comparison. Across the other level comparisons, F1 score for Game Score is lower but not significantly different to fNIRS. However, F1 scores for fNIRS & Game Score were consistently superior to Game Score across each of the three comparisons.

To explore the effects that the different measures (fNIRS, Game Score, fNIRS and Game Score) and level comparisons (E v H, E v I, H v I) had on the two classification techniques (LDA and SVM), the Estimated Marginal Means of the LDA classification was plotted. The results of the LDA plot can be seen below in Figure 44.

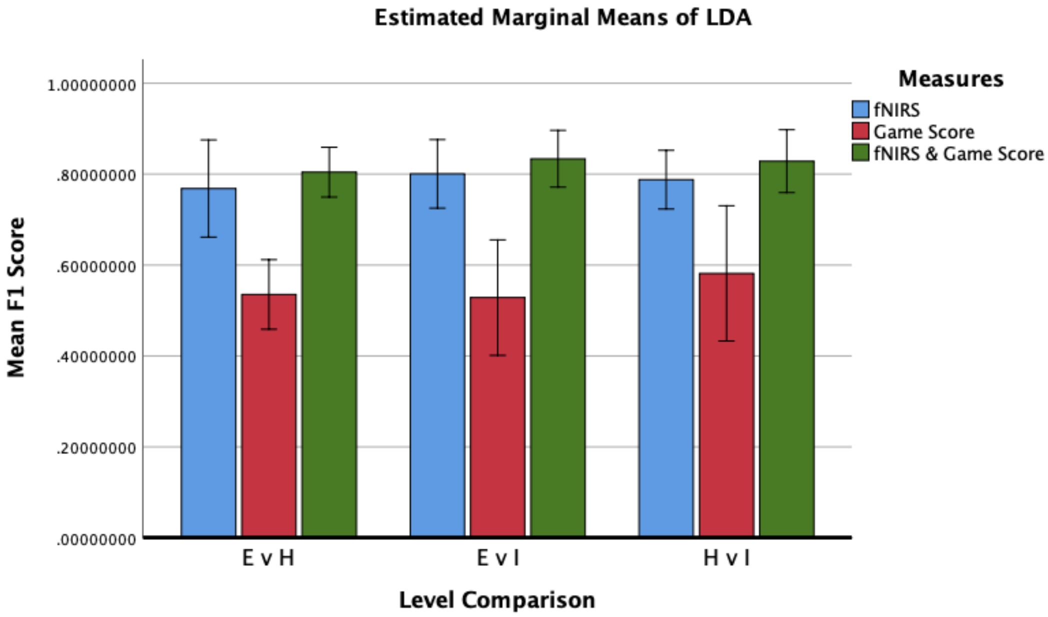


Figure 44 - Mean F1 Score for LDA Classification across Measures and Level Comparison

The main effect for Measures (being that Game alone produced the lowest F1 score overall) was only consistent across each level of comparison for the LDA model. From this result, we can infer that the Game Score measure was particularly difficutly to classify when we used a simple statsitcal model (LDA) as opposed to the more complex approach which is used for SVM classification. It is apparent when considering Figure 44 that Game Score has a considerably lower classification accuracy when compared to fNIRS and fNIRS & Game Score regardless of difficulty comparison. Paired sample t-tests of this data revealed that both fNIRS and fNIRS & Game Score provided significantly more accurate classifications than Game Score alone ($p =< 0.1$).

We also explored the effects of measures and level comparisons on an SVM classification. This plot can be seen below in Figure 45.

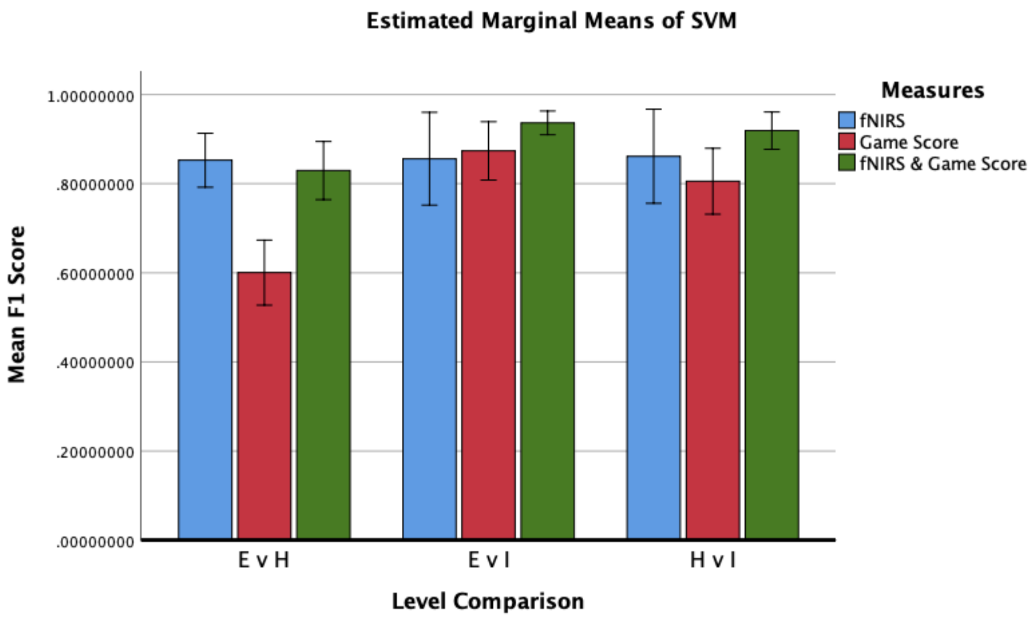


Figure 45 - Mean F1 Score for SVM Classification across Measures and Level Comparison

The paired sample t-tests used to compare the effects on SVM revealed that there was no significant difference between Game Score and fNIRS at either E v I or H v I ($p > .05$), however, there was a significant difference for the E v H comparison ($p <.01$). This result indicates that there was only a significant difference between the accuracy of fNIRS and Game Score when the game was at the lower levels. This is not surprising, as we expect that the Game Score measure is far less sensitive at the Easy level (due to a lack of change in the score resulting from less user collisions), and that differences between the lowest and middle level would only be found when using a more sensitive measure, such as fNIRS. fNIRS and Game Score being comparable in this case further proves that fNIRS is only stronger than Game Score when the comparison between levels is small.

Conversely, the difference in accuracy between Game Score and fNIRS & Game Score was always significant when an SVM was used, regardless of the level of difficulty ($p = <.02$). Again, this is not a surprising result, as our classification results alone indicate that a combination of fNIRS and Game Score data is always superior to Game Data alone.

The results presented in this chapter (specifically regarding the paired sample t-tests which were carried out) indicated that a combination of fNIRS & Game Score will provide a more accurate level of classification than Game Score alone, that an SVM classification is more effective than an LDA classification, and that the most difficult level comparison to classify is E v H.

We can use the results obtained from this study to inform the choices made for Study 3, especially regarding the best classification technique to use for the data types that are collected.

4.4 Discussion

We expected that the use of a fNIRS montage, which recorded data from two distinct cortical networks (proven in the literature to have a working association), would provide a classification rate that would be an improvement on previous fNIRS classifications. We also expected to see an increase in activity in the dorsal network as game difficulty increased. Subjectively, we expected that the three created levels of Ribbon Rush would provide a distinctly different experience to each player and that this difference would provide a statistically significant difference on pain tolerance between each of the three levels.

Our hypothesis that an adapted fNIRS montage would provide accurate classification rates was supported with classification accuracy rates between 86.4% - 91.1% were found between all three possible game conditions. These classifications resulted from a connectivity analysis, which specifically focused on determining the correlations between different channels on the fNIRS montage. We hypothesised that a montage that considered two separate brain networks (dorsal and ventral) may provide more accurate classification rates, rather than focusing on one separate network, especially when considering the use of connectivity features. Classifications were consistently more accurate as the difference between levels increased. For example, higher levels of accuracy were found between Easy v Impossible in comparison to Easy v Hard (Easy v Impossible = 94.88%, Easy v Hard = 85.38%). Interestingly, classification accuracy was more consistent between Easy v Impossible and Hard v Impossible, which indicates that the classification of Easy v Hard may have been due to individual differences within the participants perception of what constitutes an Easy or Hard game, rather than because of a lack of variation between the levels of game difficulty. Overall, it appears as though an SVM classification would provide the most accurate results in terms of data classification.

Table 9 indicates that 11 HbO channel pairs were selected by the RELIEFF algorithm. Of these, 6 pairs related to correlations between the ventral and dorsal networks, whereas only 5 related to inter-network activity. Equally, Channel 6 features in 5 of the 11 channel pairs were chosen by the RELIEFF algorithm. Channel 6 is the TPJ, responsible in part for the re-orienting of attention between tasks. The analysis of selected features alone indicates that interactions occurring between Channel 6 and all other channels is highly relevant for feature selection. When analysing Figures *34-38*, it becomes clear to see that there is a pattern of activity occurring between Channel 6 and all other channels when considering game demand. Between each of the FEFs (Channels 1 and 2) and Channel 6 (TPJ), there is an increase in correlational activity as the game enters the Hard stage, and a decrease in correlational activity as the game becomes impossible, which could indicate that during the Hard level (where we expect the game to provide the most distracting experience), there is more activity relating to ensuring that only task-relevant visual stimuli are considered by the participant. At Channels 4 (Left IPs) and 5 (Right IPs), a decrease in correlational activity is found when considering Channel 6 (TPJ), as the game becomes more difficult. This decrease is linear, i.e. there is a reduction between Easy and Hard, and a further reduction between Hard and Impossible, which indicates less correlation activity at the gradual increase of difficulty. This pattern can be observed in both Channels 4 and 5, which indicates that ventral activation may be reduced as the game becomes more difficult. This could indicate that filtering is occurring within the brain, which is preventing non-task related stimuli from becoming salient to the participant.

There was a statistically significant difference between the TLX scores between all three levels of game difficulty. This indicates that participants did find each level to be significantly different from each other. The results of the Motivation questionnaire indicate that there was a statistically significant difference in participants motivation levels between the Easy and Hard condition, and the Easy and Impossible condition, but not between the Hard and Impossible condition. This indicates that the hardest levels to distinguish between, in the participants opinion, was Hard v Impossible. However, classification rates for Hard v Impossible were more accurate than Easy v Hard – which could indicate that the game score data was affecting this classification more significantly. The results of the TLX and Motivation questionnaires indicate that there was a significant difference between the three game conditions. From observing the data relating to CPT submersion times, there is a noticeable difference between pain tolerance for each of the three levels of game difficulty. However, this difference was not statistically significant. There was, however, a significant difference between pain tolerance when a game distraction was used, compared to no game distraction.

The study detailed in this chapter has provided some improved results in comparison to previous research around the topic of fNIRS classification: Classification rates in the Verdier (Verdière et al., 2018) study were 80.5% a result that our study has improved upon. This could be due to the selection of targeted cortical networks that delivered accurate classification of game demand due to the capturing of association between distinct sites. Overall, both the results from feature selection and the results displaying correlational HbO activity between the dorsal and ventral networks indicate that the Corbetta (2008) (Corbetta et al., 2008) model was the correct choice for monitoring participants perception of game difficulty. It is clear from the interactions between the Dorsal network and the TPJ in particular that there is reduction in activation towards re-orienting activity, and an increase in activity, which indicates non-salient stimuli being filtered from the participants field of view and focus.

There are a number of limitations to this study. The reduction in classification accuracy, which was found between the Easy and Hard levels of gameplay, could be due to a lack of variation between the two levels of game difficulty. The only modifying factor for game difficulty was the overall speed of the game. Although, to a novice player, even reducing the speed of the game may not be enough to make the game easier for them. It is possible that variation of NPCs on the track could create a stronger distinction between difficulty levels, particularly for novice players. Another limitation of this study is the reliance on the 10-20 system for optode placement. Although the areas of the brain are relationally consistent between participants, the size and shape of the participants head can mean that the 10-20 placement is not always accurate. A more accurate way of calculating optimal optode placement positions is to use tools, such as a Polhemus digitiser (Artinis Systems), to determine the exact position of the desired area of the brain for each participant. Using equipment, such as this, may improve data collection, and therefore also improve data classification. We also believe that there may still have been a high level of inter-participant variability inherent in this study. Unfortunately, the nature of the CPT means that a participant cannot set their pain threshold baseline, which can lead to inter-participant variation. This variation could be the cause of the lack of statistical significance. In future studies, it may be useful to consider overall variation in pain tolerance and exclude participants whose data could be considered an outlier from the entire dataset. Reducing the inter-participant variation may provide a clearer indication of whether statistical significance could be found. It is imperative when creating a game with the intention of distracting from pain that inter-participant variation does not skew the results either positively or negatively.

Moving forward, the main focus of our work will relate to the creation of a Neuroadaptive game. Determining what should be considered an acceptable classification rate for a neuroadaptive game is difficult, as there is no preceding literature relating to the development of a neuroadaptive game for pain management. However, classification accuracies above 80% are generally considered acceptable in alternate fNIRS studies (Naseer & Hong, 2015). Considering the results that have been gathered in this study, we are satisfied that the classification results we have managed to gather will be sufficient, at least in the first instance of producing and testing a neuroadaptive game. For the combined classification of fNIRS and Game Score data, the Easy and Hard v Impossible conditions had an error rate of around 5%, which means that we should expect to see errors in classification in roughly every 200 seconds of gameplay.

It is also important to consider the relationship between game difficulty and pain tolerance, particularly in relation to how this may be reflected in the Corbetta (2008) (Corbetta et al., 2008) model. It is expected that a higher level of distraction, caused by the game difficulty, would be reflected within the CPT results; however, this was not found. It would be especially interesting, therefore, to analyse whether there was any correlation between the ventral and dorsal networks, which reflect the difference between game difficulty level and pain tolerance (if this did indeed occur). Not only would this provide us with a clearer idea as to whether the TPJ is re-orienting attention when there is pain stimuli present, but it may provide us with a clearer idea as to whether one level of game difficulty is more efficient at distracting from pain and provide us with results, which may have been obfuscated in the CPT results due to inter-participant variability. Although it would be interesting to explore this avenue, we do not see it as particularly relevant for the creation of a neuroadaptive game. We plan to develop the neuroadaptive game based on the classification of difficulty levels alone. However, it would be interesting, following the creation of a neuroadaptive game based on game demand, to also create a neuroadaptive game based on pain responses, specially to understand whether either version would provide a more accurate model for pain distraction. Moving forward, we plan to use the dataset that was gathered in Study 2 to inform the creation of a neuroadaptive game. The classification results reported in this study are from binary classifications (despite there being three levels of game difficulty), which indicates that continuing to use binary classification during the development of a neuroadaptive game may be the most efficient way forward.

Chapter 5 - Design and Development of a Real-Time Neuroadaptive Game

5.1 Introduction

An adaptive game is any type of game wherein dynamic difficulty adjustment (DDA) is used. DDA makes real-time adjustments to a gameplay scenario in an attempt to engage a player, and to maintain this engagement (Zohaib 2018). This is achieved by preventing undesirable states, such as boredom or overload, and encouraging a state of flow or immersion. DDA has been studied heavily for the purpose of achieving positive cognitive states (Hunicke, 2005; Jennings-Teats et al., 2010; C. Liu, Agrawal, Sarkar, & Chen, 2009), and recently games companies, such as EA (City, 2017)[8], have begun to use these algorithms in their commercially available software. DDA can be utilised to maximise the level of game engagement in order to distract from pain. It is already understood that games that can adapt to match the skill level of the player can be effective in reducing the experience of pain (Weiss 2011).

Although DDA can be applied via the use of game metrics and physiological measures, a neuroadaptive game, wherein the adaptive controller is influenced by brain activity (Zander et al., 2016), could potentially be even more effective in encouraging the participant to enter the flow state. As such, it is particularly important for our research that we can create an effective distraction technique through a neuroadaptive video game. Neuroadaptive games should be more sensitive to changes in player engagement and attention, compared to performance outcomes, because neurophysiological data itself is sensitive and continuous. The neuroadaptive game created in this chapter works on a negative control loop (Pope et al., 1995), wherein the demand of the game is designed to increase if a player is disengaged, due to boredom, and decrease if the player's disengagement is due to excessive demand (Ewing et al., 2016). We hypothesise that a neuroadaptive game will be more effective than a standard game at reducing the experience of pain.

This chapter will discuss the materials and methods that were used for the creation of a neuroadaptive game, as well as the rationale behind the implementation of the specific methods that were used.

5.2 Communication between Hardware and Software

Chapter 2 discussed the different hardware and software components that were used throughout the course of the studies. For the real time study, the same hardware was used, and software techniques were employed in order to set up communication between this hardware. These techniques are discussed in detail below.

[8] Patent US20170259177A1

5.2.1 Real-Time System Architecture

The hardware configuration that was used for the creation and use of the neuroadaptive game is illustrated below in Figure 46.

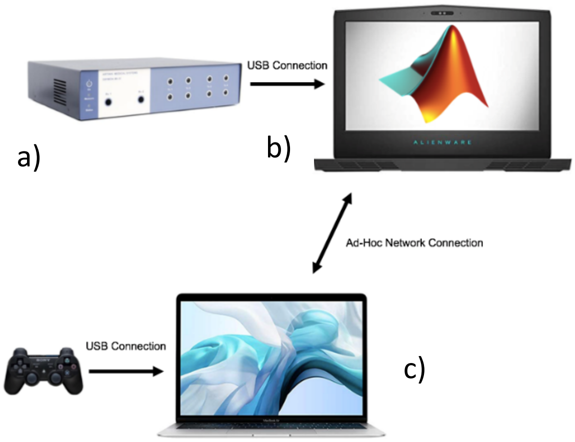


Figure 46 - System architecture of the neuroadaptive system that was used for the creation and application of the neuroadaptive game, including a) Oxymon Mk III fNIRS system, b) an Alienware 15 R4 Laptop Computer and c) a 13-inch MacBook Air

The system was composed of standard hardware components. An Oxymon Mk III (see Figure *46* (a)) was used to collect fNIRS data from the participants. This system was comprised of two control boxes, which were connected to an fNIRS cap. The cap was worn on the participants head for the duration of game play, in order to collect neurophysiological data. A 15-inch Alienware 15-R4 laptop computer, with an 8th generation Intel Core i7 processor (see Figure *46* (b)), was used for data collection, pre-processing and classification. Oxysoft, a proprietary software created by Artinis Medical Systems, was used to display the fNIRS signal to the experimenter, whilst MATLAB (v R2018b) was used to stream the data in real-time. Real-time streaming was achieved via a 10-second buffer, which gathered data from port 1972 using the localhost command. A 13-inch MacBook Air (Early 2015), with a 2.2 GHz Intel Core i7 processor (see Figure *46* (b)), was used to deploy the game Ribbon Rush (as described in Chapter 2). The player's current score was then outputted to a .txt file every 2 seconds, and was also continually displayed to the player on-screen. In order for the Game Score information, which was recorded by Ribbon Rush on the MacBook, to be transferred to the Alienware laptop, an ad-hoc network connection was established between these two computers.

5.2.2 Oxymon to Alienware

The Oxymon system was connected to the Alienware laptop via a USB 2.0 to USB 2.0 connection. Oxysoft was installed on the Alienware laptop and used to display the fNIRS signal to the experimenter. A MATLAB session was also running at the same time as the Oxysoft data collection session and was used to stream the data from Oxysoft in real-time. Real-Time streaming was achieved via the use of a buffer, which gathered data from port 1972 using the localhost command. The localhost command is used to gather information that is being streamed from Oxysoft, to port 1927, and then the live signal is converted (via the buffer) to a matrix of data. The localhost command can only be executed if the data is being sent and received via two different pieces of software running on the same computer. Due to the classification of data being performed over 10 seconds of data at a time, data was gathered from the buffer (which can be considered as the location where streamed data is held) in 10 second blocks.

5.2.3 MacBook to Alienware

In order for the Game Score information (which was recorded by Ribbon Rush on a 2015 MacBook Air with macOS High Sierra v10.13.6) to be transferred to the Alienware laptop, an ad-hoc network connection was established between these two computers. Ad-hoc refers to a makeshift or improvised wireless network that does not require alternative hardware (other than two or more WiFi-enabled devices) to be established. An Ad-Hoc network is a temporary network, and, once this network is established, any files or folders that are placed within a shared folder within this network can be seen by any device that is connected to this network. Ad-Hoc networks do not require additional hardware (such as an access point or Wi-Fi router) and are easy to establish, with instantaneous transfer speeds.

The Ad-Hoc network was set up between the MacBook and Alienware laptops, and the game Ribbon Rush was played from a .exe file stored within the shared folder on this Ad-Hoc network. This meant that any files contained within the folder could be accessed and edited by either computer. The path of this shared folder was: `\\kellyannsair.mshome.net\RibbonRush_Dec2018\`and the file where the game score data, which was transferred from the MacBook to the Alienware computer, was stored was called `fromRibbonRush.txt`.

5.2.4 Alienware to MacBook

The same Ad-Hoc network that was created for transferring data from the MacBook to the Alienware laptop was used to transfer data from the MacBook to the Alienware computer. Again, the path of the shared folder was: `\\kellyannsair.mshome.net\RibbonRush_Dec2018\` and the file in which the current difficulty level, which was transferred from the Alienware computer to the MacBook, was called `toRibbonRush.txt`.

5.3 Real-Time Biocybernetic Loop

The neuroadaptive game relies on a biocybernetic loop for the purpose of making informed changes to the difficulty of the game, based on neurophysiological data. The system is built upon a closed loop control, wherein data is collected from a participant and decisions are made based purely on that data, without input from either the participant or the experimenter. All of the potential changes, which could be made by the biocybernetic loop, and the conditions that must be fulfilled to make these changes, are imbedded within the neuroadaptive system. It must therefore be designed with considerations to every possible outcome and what should be performed following that outcome. The loop that has been designed in this chapter collects, pre-processes and classifies both fNIRS and game score data. This classification is then used to make a decision about the direction and strength of the game difficulty to change. This process has been designed to encourage and maintain a state of engagement and to reduce boredom or overload.

5.3.1 fNIRS Data Collection and Pre-processing

fNIRS data was collected from the brain networks that are associated with the ventral and dorsal networks (See Chapter 4). Figure 47 is a diagram of the montage that was used for fNIRS data collection:

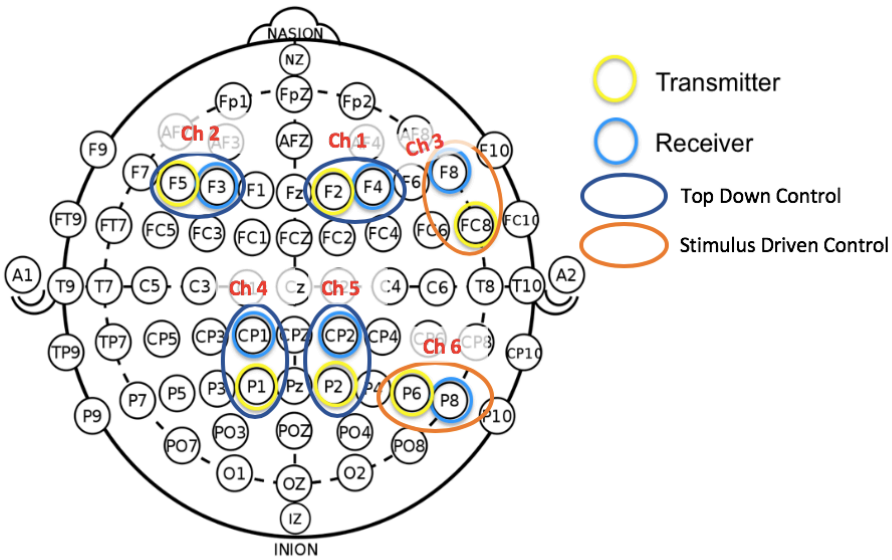


Figure 47 - fNIRS montage used for the creation of the Neuroadaptive game

Once data had been collected, but before classification of fNIRS data could be carried out, pre-processing of the fNIRS data was required. The pre-processing protocol is designed to remove noise from the fNIRS signal, which is not directly related to brain activity, but may still have an effect of the classification accuracy. The purpose of this specific pre-processing protocol is to ensure that the most accurate classification can be made. A description of this protocol has been presented within section 4.2.5 Signal Processing Pipeline and in Figure 8 (Chapter 4), which illustrates the pre-processing protocol that has been implemented during the real-time stage.

5.3.1.1 Game Score Data Pre-processing

Following the creation of features from the fNIRS dataset, features were also created from the game score data. These features were Mean (the average game score), Peak (the maximum game score) and Variance (the average of the squared differences from the mean.) Game score data was recorded every 2 seconds, which provided five data points over a 10 second epoch. Calculating the Mean, Peak and Variance of these 5 data points provided one data point for each of the features over a ten second period. Game score features were combined with the 19 fNIRS features for classification.

5.3.1.2 Data Classification

Typically, once features have been created from data, feature selection methods are used to determine a subset of important features for classification. Feature selection is used to reduce the size of a dataset, whilst increasing classification accuracy. In this case, relevant features have already been determined prior to the creation of the neuroadaptive game (Chapter 4), and so only these features were created. Feature selection was carried out in this way to reduce the overall computation time of the real-time classification process and to ensure that accurate classifications could be carried out in a real-time fashion.

An illustration of the neuroadaptive controller can be seen below in Figure *48*. Once the two feature sets (fNIRS and game score) had been combined, the data were ready for classification. It had already been determined via the results of Study 2, which can be seen in Chapter 4, that the most accurate classification results could be achieved when a binary classification method was applied, rather than a multiclass classification. It was also pre-determined that a Support Vector Machine (SVM) would provide the highest level of classification accuracy for this particular dataset. For this reason, two individual SVMs were created. The determination of which SVM should be used was based on the current difficulty level for the game. The training data that was used for the real-time classification was the data that was collected during Study 2. This data was collected from three categories of game difficulty: Easy, Hard and Impossible. For this reason, two bands of game difficultly were created for the determination of which SVM should be used. If the current difficultly of the game fell within the Easy/Hard band, which was a game speed between 30 and 70, then the Easy/Hard SVM was used. This contained training data from the pre-defined Easy and Hard levels. Similarly, if the current difficulty of the game fell within the Hard/Impossible band, which was a game speed between 71 and 150, then the Hard/Impossible SVM was used.

Real-Time adaptation in the context of fNIRS data can be considered to be ambiguous – there is roughly a 7 second delay present in fNIRS data between the onset of stimuli and the representation of response to this stimulus in the brain. For this reason, when creating a 'real-time' system using fNIRS data, we must consider how effective a truly 'real-time' system would be. Although we have created a real-time fNIRS filtering protocol, for the sake of fNIRS classification the protocol is not truly working with real-time data to the second – there are multiple reasons for this. Firstly, although the protocol which we designed could have worked on a second-by-second basis, we did not feel as though classifying a second of fNIRS data (which would have only been 10 samples) would provide us with an accurate classification. For the sake of providing an accurate and informative classification, we decided to collect and classify 10 seconds of fNIRS data at a time. Along with providing more information to create an accurate classification, extending the fNIRS classification window to 10 seconds also accounts for haemodynamic delay that is prevalent in fNIRS data. Hemodynamic delay means that the BOLD response to stimuli does not occur for up to 7 seconds after the onset of the stimulus. Allowing for a 10 second delay before beginning to filter and classify data ensures that the response to changes in game difficulty are captured for classification. It is important to ensure that the data which is collected for classification contains responses to the most recent game difficulty changes, to ensure that adjustments are not made in the incorrect direction.

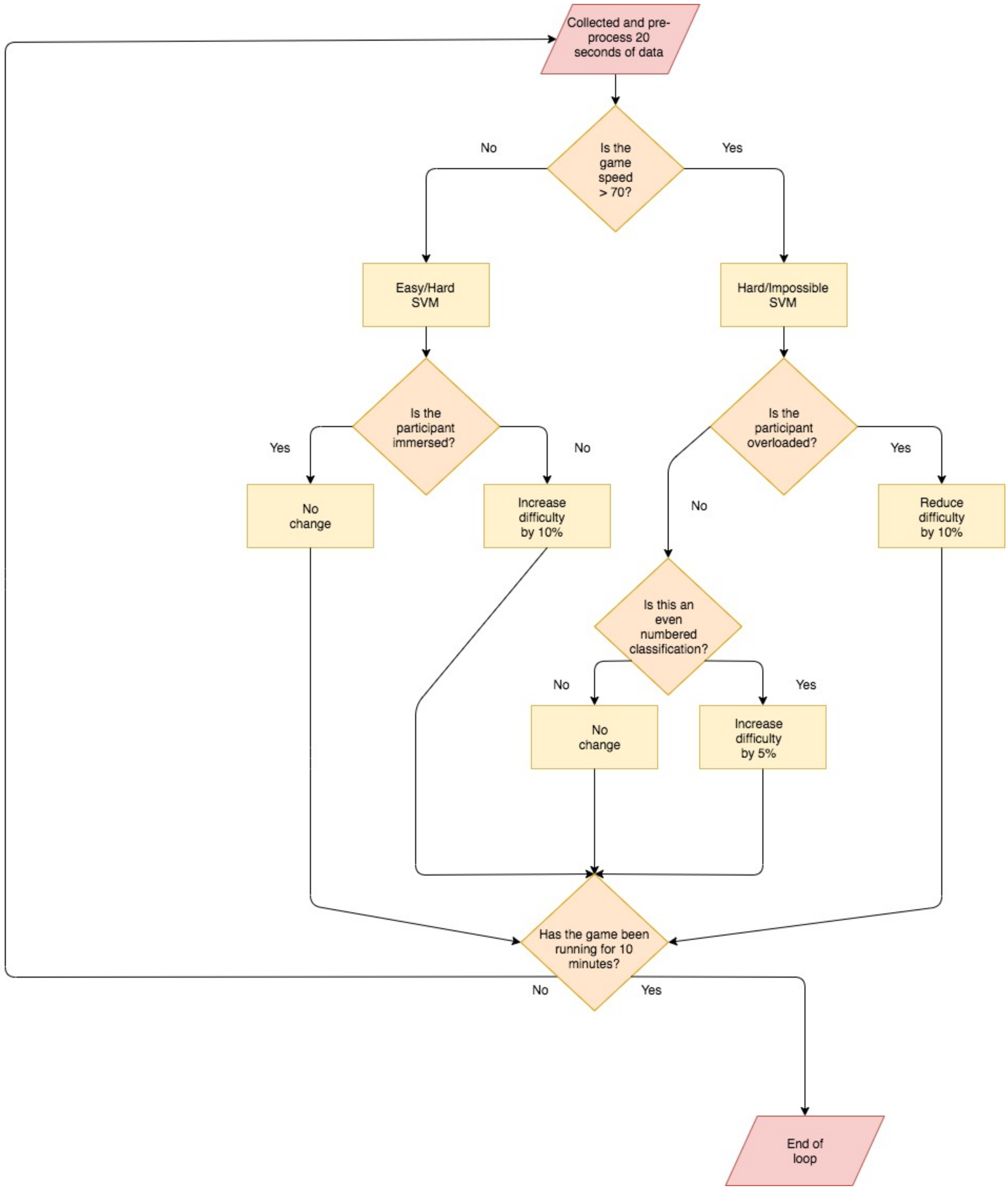


Figure 48 - Flow diagram illustrating the data collection and classification process which was followed for the neuroadaptive game

One of the most important aspects of classifying real-time data is to determine when data collection should begin following an adaptation of game demand. This is particularly relevant when considering the hemodynamic delay, which is inherent to fNIRS data. It is important to ensure both that there is sufficient time between each adaptation to ensure that the physiological response to such an adaptation can be measured. Equally, it is important to ensure that the time period between each adaptation isn't too long, as an incorrect adaptation could be made, which would not produce the desired effects of the adaptive game. For this reason, it was decided that a period of 30 seconds would be established between a change being made to the difficulty of the game, and data collection for the next classification period beginning.

It is worth noting that, for this study, we chose to use subject-independent (SI) (or universal) classification. This means that the classifier was trained on a dataset of 20 participants, and all of this data was later used for the real-time classification of each participant. Using subject independent classification provided us with a larger, richer dataset to train our classifier – however, there are downsides to subject independent classification. The biggest drawback of SI classification is that the data is not specifically tailored to each participant – meaning that there are individual differences present in the data which may have affected classification results. Using a large dataset should, in theory, remove the reliance on individual differences and shouldn't affect the classification significantly, but there is still a possibility that this may have happened. An alternative to SI classification is Subject Dependant (SD) classification – where the classifier is trained using a participant's data, and the classifications are made on that same participants real-time data. A SD classification can yield more accurate and personal results. There are, however, also drawbacks to a SD method. The biggest drawback in this case is that we would have first had to gather data from each participant at each level of difficulty both with and without pain in order to train our classifier. Although this is possible and viable in some scenarios, thinking of the 'bigger picture' of potentially using the neuroadaptive game in a clinical setting, we felt it would not be valid to require this level of data collection prior to the use of the neuroadaptive game, and therefore the start of a clinical procedure.

5.3.1.3 Functions for Communications between Hardware Components

In order for the data to be transferred between Oxysoft and MATLAB, this data was pushed into a buffer by Oxysoft, and then received from the buffer into MATLAB. An example of the code that was used for this purpose can be seen below in Algorithm 4. Once 10 seconds of fNIRS data had been collected, the data were pre-processed and classified. There are three possible classifications that can be made – *Not Immersed (E_C), Immersed (H_C),* or *Overloaded (I_C*). These three letters referred to the Easy, Hard and Impossible labels that were applied to the training data. Depending on the result of the classification, one of these variables (either *E_C, H_C* or *I_C*) would be assigned a value of 1, whereas the value of the two alternate classifications would be assigned a value of 0. In order for the real-time adaptation algorithm to determine the action that should be taken depending on the classification result, an if-else loop was established. This loop, which can be seen below in algorithm 4 determined the immersion level of the participant.

Algorithm 4. Determination of Immersion

```
if E_C > H_C
    immersion = 1;
    disp('Score indicates player not immersed');
else
```

```
    if E_C < H_C && H_C > I_C
        immersion = 2;
        disp('Score indicates player immersed');
    else
        if I_C > H_C
            immersion = 3;
            disp('Score indicates player overloaded');
        end
    end
end
if immersion == 1 % Not immersed
    level_Controler = current_Level + 10;
    fileID_Level = fopen...
        ('\\kellyanns-
air.mshome.net\RibbonRush_Dec2018\toRibbonRush.txt','w');
    fprintf(fileID_Level, '%g\n',level_Controler);
    fclose(fileID_Level);
    disp('Level increased');
else
    if immersion == 2 && current_Level <= EH
        % Do nothing
    else
        if immersion == 2 && current_Level > EH % Immersed in hard game
            alternate_Test = mod(alternate,2);
            if alternate_Test == 1
                level_Controler = current_Level + 5;
                fileID_Level = fopen...
                    ('\\kellyanns-
air.mshome.net\RibbonRush_Dec2018\toRibbonRush.txt','w');
                fprintf(fileID_Level, '%g\n',level_Controler);
                fclose(fileID_Level);
                disp('Level increased');
            end
        else
            if immersion == 3 % Overload
                level_Controler = current_Level - 10;
                fileID_Level = fopen...
                    ('\\kellyanns-
air.mshome.net\RibbonRush_Dec2018\toRibbonRush.txt','w');
                fprintf(fileID_Level, '%g\n',level_Controler);
                fclose(fileID_Level);
                disp('Level decreased');
```

```
            end
        end
    end
end
```

- Line 1 instantiates a nested if loop, which serves to determine the user's level of immersion and alert the experimenter of the result. Line 1 determines whether there are more easy classification indicators (*E_C*) than there are hard classification indicators (*H_C*).
- Line 2 sets the value of immersion to 1, if the classification indicates that the player is not immersed.
- Line 3 alerts the experimenter that the player is not immersed.
- Line 4 instantiates the first 'else' statement, which is used if the conditions of the first if statement (line 1) are not satisfied.
- Line 5 is an if statement which states that, if there are less easy classification indicators (*E_C*) than hard classification indicators (*H_C*) and less hard classification indicators (*H_C*) than impossible classification indicators (*I_C*) then the player is immersed.
- Line 6 sets the value of immersion to 2, if the classification indicates that the player is immersed.
- Line 7 alerts the experimenter that the player is immersed.
- Line 8 instantiates the second 'else' statement, which is used if the conditions of both the first and second if statements are not satisfied.
- Line 9 is an if statement which states that, if there are more impossible classification indicators (*I_C*) than there are hard classification indicators (*H_C*) then the player is overloaded.
- Line 10 sets the value of immersion to 3, if the classification indicates that the player is not immersed.
- Line 11 alerts the experimenter that the player is overloaded.
- Lines 12-14 end the series of nested if loops.
- Line 15 beings a second series of nested if loops. Each of these loops can be used, depending on the value of Immersion which was previously determined by the first series of nested if loops. Line 15 determines whether the value of immersion is equal to 1 (not immersed).
- Line 16 increases the difficulty level (*level_Controller*) by finding the current difficulty level (*current_Level*) and adding 10 to this value.
- Lines 17-19 gain access to the file stored on the MacBook which determines the current level of the game, and gives the algorithm write access ('w') to edit this file.
- Line 20 prints the new level value to the difficulty level file, to adjust the difficulty of the game.
- Line 21 closes the file to prevent further editing.
- Line 22 alerts the experimenter that the difficulty level has been increased.
- Line 23 defines the first else statement, which is used if the conditions of the first if statement are not satisfied.
- Line 24 defines the second nested if statement, which determines whether the immersion level is set to 2 (participant immersed) and if the current level is easy. If this is the case, no changes are required.

- Line 25 is a comment to state that, if the conditions of line 24 are met, then the algorithm does not need to perform any actions at this time.
- Line 26 defines the second else statement, which is used if the conditions of any previous if statement are not satisfied.
- Line 27 determines whether the participant is immersed, and the current level of game difficulty.
- Line 28 instantiates a check, which determines when the last change to game difficulty was made. During the hard level, every alternate pass of the loop increases the difficulty of the game even if the participant is immersed, to ensure that the participant does not become bored.
- Line 29 is an if statement, which determines if the loop is on its first or second pass since the last difficulty adjustment was made. If the loop is on its second pass, then the difficulty of the game is increased.
- Lines 30-37 determine the current level of game difficulty, adjust the level of difficulty upwards by 5, edit the difficulty of the game in the file stored on the MacBook and close the file. Line 36 alerts the experimenter that the level of difficulty has been increased, and line 37 ends the loop.
- Line 38 defines the third else statement, which is used if the conditions of any previous if statement are not satisfied.
- Line 39 is an if statement, which determines whether the player is overloaded (immersion = 3). The purpose was to determine the level of game difficulty the classifier assumed the player was experiencing, based on their game score and fNIRS data. The variable 'immersion' could either be set to 1 (not immersed), 2 (immersed) or 3 (overloaded).
- Lines 40-46 decrease the level of difficulty by 10 (a number found to create a subtle difference in the gameplay experience) , and access and change the file on the MacBook which determines the level of difficulty which the game should be set to.
- Lines 47-50 close all remaining open if and else statements of the algorithm.

Following the determination of the immersion value, Lines 1-50 display a further if-else loop which was established to control the potential change in game difficulty. This loop was established to determine the action that should be taken depending on the classification result, and then to carry out this action. The variable '*current_Level*' refers to the current level of game difficulty, and the variable '*level_Controller*' refers to the new level, which should be applied to the game based on the classification. In order to manipulate the difficulty of the game, the current difficulty of the game was gathered from the MacBook using MATLAB. Changes were made to the current level of difficulty, via further MATLAB code. Difficulty could also be increased by 5%, decreased by 10% or remain unchanged, depending on the classification outcome.

5.3.2 Pilot Study

At this point, testing of the developed system was focused on the actual functions of the game, rather than testing its efficacy as a distraction from pain. As the game is neuroadaptive, the actual behaviour (in terms of when the difficulty will adjust, and in which direction) cannot be estimated. As such, the changes that are expected are an unknown quantity. However, it is expected that each participant will have a different experience. To test this functionality, a pilot test of five participants was carried out. These participants played for ten minutes, and the behaviour of the game was recorded to ensure that the neuroadaptive element was performing differently for each participant. Figure 49 below shows the average game performance information for each participant:

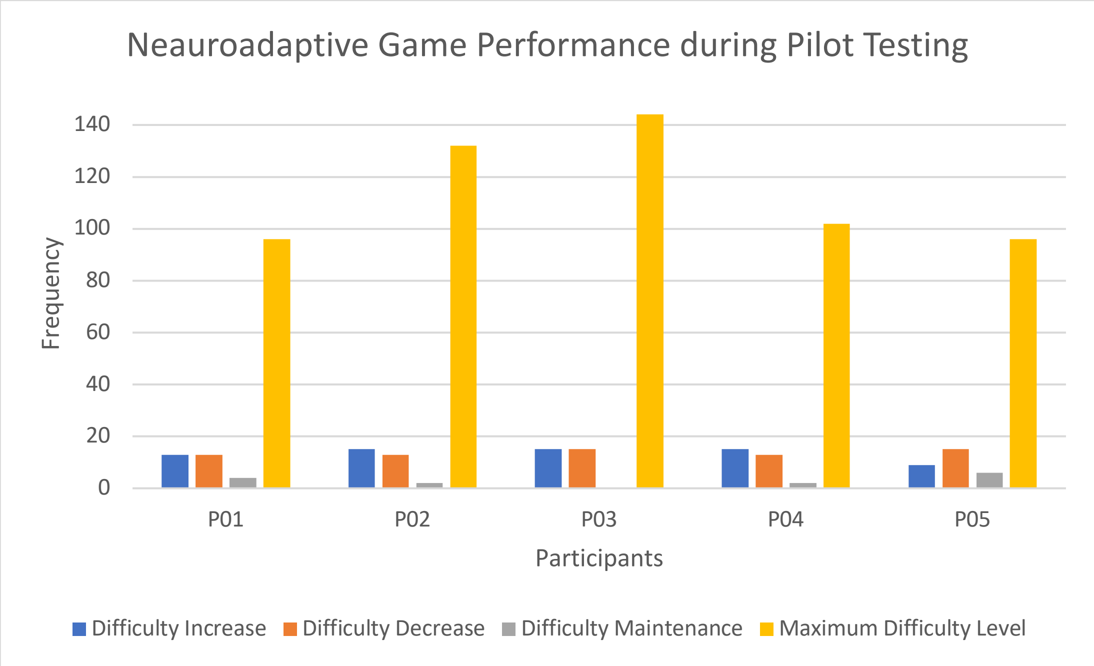


Figure 49 - Figure depicting difficulty alteration data, which was gathered during pilot testing

It can be inferred from Figure 49 that the neuroadaptive game is functioning as expected in terms of providing a different gaming experience for each participant. The maximum difficulty level (yellow) indicates that each participant had a different maximum difficulty level, i.e. the point at which the game became too difficult for the player. Although the frequencies of difficulty increase/decrease/maintenance are less obviously different, there is still a definable difference between these changes for each participant. These results indicate that, although we do not yet know whether the game is functional in distracting a participant from pain, it is definitely providing a different gameplay experience for each individual participant.

5.4 Summary

This chapter has outlined the steps that are required to measure and adjust immersion in real time. The first steps in measuring immersion is to collect and process data. The processes which have been used to process data in real time are outlined and explained in Algorithms 1-3. Following data processing, data is classified in real time to determine the participants current level of immersion. Algorithm 4 is used to categorise the immersion classification and adjust the level of game difficulty, if this is required. Following the creation and implementation of these algorithms in real time, a pilot study was carried out to ensure that the algorithms and experimental protocol were functioning as expected.

We were satisfied that the pilot study produced results in line with what we had intended and expected, indicating that the real-time pipeline was suitable for deployment in a full study. This study will be discussed in more detail in Chapter 6.

Chapter 6 - Study 3: Evaluation of the Developed Neuroadaptive Game

6.1 Introduction

Physiological computing refers to any computer system wherein the system is influenced by physiological data collected from a human in real-time, i.e. psychophysiological or neurophysiological data (S. H. Fairclough, 2009b). Physiological systems will use or display this collected data, and all such systems rely on a biocybernetic loop in order to collect, process, analyse and translate data into a format that can be utilised. Biocybernetic loops (Pope et al., 1995) define how and when data is collected and used by a physiological computing system. Such loops can be classified into one of two categories: open or closed loops. Although both types of loops require data to be gathered from a human participant, an open loop system also requires a person (be it a participant or an experimenter) to make a decision based on the data that is presented. This decision is subsequently used to inform the biocybernetic loop of which adjustments are required to the system. In a close loop control, data is collected from a participant and decisions are made based purely on that data and require no further input or decisions to be made. All of the closed loop decisions are imbedded within the design of the biocybernetic loop[9].

Until 1995, biocybernetic loops were used only to reflect physiological data. At this time, Pope et al (Pope et al., 1995) published a paper that introduced what we now refer to as the biocybernetic loop. This loop, which concerned biocybernetic adaption, introduced the idea of using physiological data to inform a system of changes that were required to how the system operated. Pope et al (Pope et al., 1995) collected Electroencephalogram (EEG) data from participants with the intention of determining their attentional state. The goal of the study was to investigate whether real-time adaptation to the control of a flight simulator system could alter the attentional states of the pilots who were using it. This study was carried out to investigate the issue of pilot inattention due to the improvements and increased use of autopilot systems, with the intention of discovering the optimal mix of pilot control and autopilot control, which should be employed in order to ensure that the pilot is constantly maintaining a state of mental activation and awareness. Both positive and negative feedback loops were tested during this study. In a negative feedback loop, the task was switched from automatic to manual when the slope (calculated through EEG data) began to decrease. In a positive feedback loop, the task was switched to automatic when the slope decreased. In 2000, this work was expanded by Prinzel et al. (Prinzel, Freeman, Scerbo, Mikulka, & Pope, 2000), who focused on identifying the effects of moderating task load based on EEG data. This study had two conditions: high workload and low workload. Similarly to Pope et al (Pope et al., 1995), both a positive and negative feedback loop were implemented. The results of this study indicated that more changes were made (between manual and automatic control) when a negative feedback loop was employed, in keeping with the results found by Pope et al (Pope et al., 1995). This study also found that more changes in the task mode (manual or automatic) were found

[9] https://www.dataforth.com/closed-loop-vs-open-loop-control.aspx

in the high workload condition as opposed to the low workload condition. The most pertinent result found in this study, when considering biocybernetic adaptation, was that participants performed better and rated their workload as lower in a negative feedback loop system, as opposed to participants in the control group. The participants in the control group performed the same tasks as the participants in the closed-loop group, but without the benefit of the adaptive system. These results indicate that the moderation of workload and the improvement of task results can be achieved through biocybernetic adaption.

The main purpose of the studies reviewed in this section has been to bring about system adaptation, be that in the switching between manual and auto modes or the overall control of a system. It is common in gaming scenarios that changes to the system (particularly the level of difficulty) are required to provide an enjoyable and achievable game. When the difficulty of a game changes throughout gameplay (as opposed to a level of difficulty being chose prior to the start of the game), this is known as Dynamic Difficulty Adjustment (DDA) (Hunicke, 2005). The aim of real-time DDA is to provide a more enjoyable gameplay experience to both novice and experienced users. Often, DDA is based around in-game performance (for example, the players score or in-game achievements), although studies have been undertaken where DDA is informed by neurophysiological or physiological signals. Generally, these studies aim to detect and manipulate a player’s attentional state, via a physiological marker (i.e. using heart rate to infer anxiety level and using anxiety level to assume game involvement). A 2009 study carried out by Liu et al (C. Liu et al., 2009) found that 77% of participants self-reported a more challenging level of game difficulty when DDA was manipulated by physiological signals, compared to DDA that relied on game metrics only. DDA that receive physiological signals as input have been found to be affective in increasing a player’s enjoyment of the game (Bontchev, 2016).

Biocybernetic loops have often been used to determine engagement and make adjustments in order to enhance or maintain a high level of engagement (Ewing et al., 2016; S. H. Fairclough, 2010; S. H. Fairclough, Gilleade, Ewing, & Roberts, 2013b; Muñoz, Gouveia, Cameirão, & Badia, 2017). This study aimed to examine the effects of a neuroadaptive game, using a closed biocybernetic loop system, on pain tolerance and subjective mental workload. A neuroadaptive game was created with the intention of investigating the hypothesis that a neuroadaptive game (or, a game that can adapt based on neurophysiological reactions to stimuli) will be more effective at reducing the subjective experience of pain and increasing pain tolerance than a random adaptive game. This chapter will discuss the functionality of both the neuroadaptive game and the random adaptive game, which will be used for comparison. Both the neuroadaptive game and the random adaptive game will feature DDA.

6.2 Methodology

6.2.1 Design

Study 3 was a two-factor repeated measures study, consisting of two 10-minute games (neuroadaptive and random adaptive) delivered in a double-blind randomised order.

6.2.2 Participants

The study consisted of 20 (9 female) participants. Participants were aged between 19 and 35 (M = 26 , SD = 4.5). Prior to the beginning of the experiment, each participant gave informed consent and completed a demographic questionnaire to ensure that they were suitable to continue with the experiment. The procedure for the experiment and data collection protocol was approved by the Liverpool John Moores University (LJMU) University Research Ethics Committee and the experiment was conducted in accordance with the recommendations of the LJMU University Research Ethics Committee.

6.2.3 Hardware Specifications for Neuroadaptive Game

The neuroadaptive game required two computers to ensure that gameplay speeds and processing speeds were not negatively affected by each other. Using two computers ensured that the Central Processing Unit (CPU) of each computer could dedicate the maximum available Random-Access Memory (RAM) to each of the overarching tasks (gameplay and data processing/classification) to ensure that these processes could be performed efficiently. This hardware is discussed in further detail in Chapter 5.

6.2.4 Dynamic Difficulty Adjustment via a neuroadaptive game

To create a neuroadaptive game, fNIRS data was collected using an Oxymon Mk III, which recorded changes in optical density from the participant's brain. This was used as a measure of neuroactivity. The participant wore a cap consisting of 6 channels (6 transmitters and 6 receivers) on their head. The areas where channels were created are outlined below in Figure 50, according to the 10/20 system:

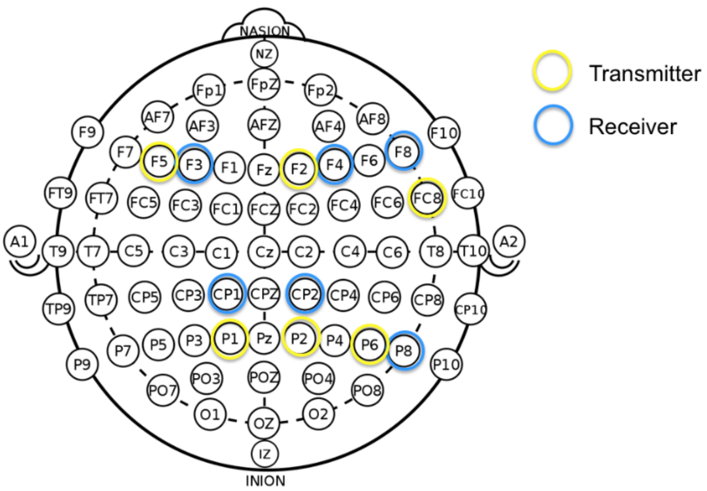


Figure 50 - Figure depicting locations of channels with reference to the 10/20 system for EEG

The score of the game was calculated based on the collision avoidance skills of the player. During gameplay, the game score increases consistently as the participant controlled the vehicle. However, if the participant collided with another vehicle, then points are deducted from their overall game score.

For example, if a participant has a score of 150 in their first 20 second period, but a score of 140 in their next 20 second period, this would indicate that the player had had at least one collision with another vehicle during their second 20 second period.

During the study, data collection and pre-processing was performed on an Alienware laptop via MATLAB, and changes to the level of game difficulty were sent via a WiFi connection to the gameplay laptop, which was a MacBook Air. The game was controlled using a Sony PlayStation 3 remote, connected to the MacBook via USB.

fNIRS data and game score data were classified in real-time via MATLAB. An SVM was used to perform this classification. Selected features from the fNIRS data and game score data were used to inform this classification. These features were established through analysis and classification of data that was collected in Chapter 4 (Study 2), which was used for training the SVM in Study 3. The same pre-processing techniques and feature creation were used in both Studies 2 and 3, to ensure that the data were comparable for classification. To ensure that the real-time processing pipeline used in Study 3 was computationally sufficient for a real time system, only the features that were determined to be significant in Study 2 were created in real time during Study 3. The real-time data processing and classification pipeline is unique to this study; to our knowledge, no other study has developed and implemented a real-time pre-processing and classification pipeline in this way.

During Study 3, a 20 second baseline period was established to gather baseline data, which was used to apply the Modified Beer Lambert Law (mBLL) to the raw optical density (OD) data collected by the Oxymon equipment. Once data had been converted from raw OD to oxygenated and deoxygenated haemoglobin, the baseline data was discarded.

fNIRS and game score data were classified at 20 second intervals. Following each classification, there was potential for the level of game difficulty to be altered. The minimum and maximum levels of the game were 30 and 150 respectively. The difference here is 120, and changes could be made to the difficultly of the game in either 5% or 10% increments, depending on which classifier was used. For example, if the game speed was 30 and the speed was increased by 10%, the new speed of the game would be 42. This was calculated using the following equations:

$$150 - 30 = 120 \tag{8}$$

$$120 / 10 = 12 \tag{9}$$

$$30 + 12 = 42 \tag{10}$$

Equation (1) calculates the difference between the upper and lower limits of game difficulty. Equation (2) calculates 10% of the sum of equation (1). Equation (3) calculates the original speed + 10% increase. If the game was in the Easy/Hard difficulty band (30-70), and classification indicated that the player was not immersed, then the difficulty of the game was increased by 10% from the overall difference between the easiest and hardest levels of the game. If classification indicated that the player was immersed, then the difficulty of the game remained unchanged.

Once the difficulty of the game entered the Hard/Impossible band (71-150), the difficulty of the game was either increased or decreased by 10%, depending on the results of the classification. If classification indicated that the player was not immersed, then the difficulty was raised by 10%. Alternatively, if the player was overloaded, then the difficulty was reduced by 10%. If the player was immersed, then difficulty would remain unchanged for one classification period. However, if the player was immersed for two consecutive classification periods, then the difficulty of the game was increased by 5%, to prevent boredom. A flow chart depicting this process can be seen below in Figure 51.

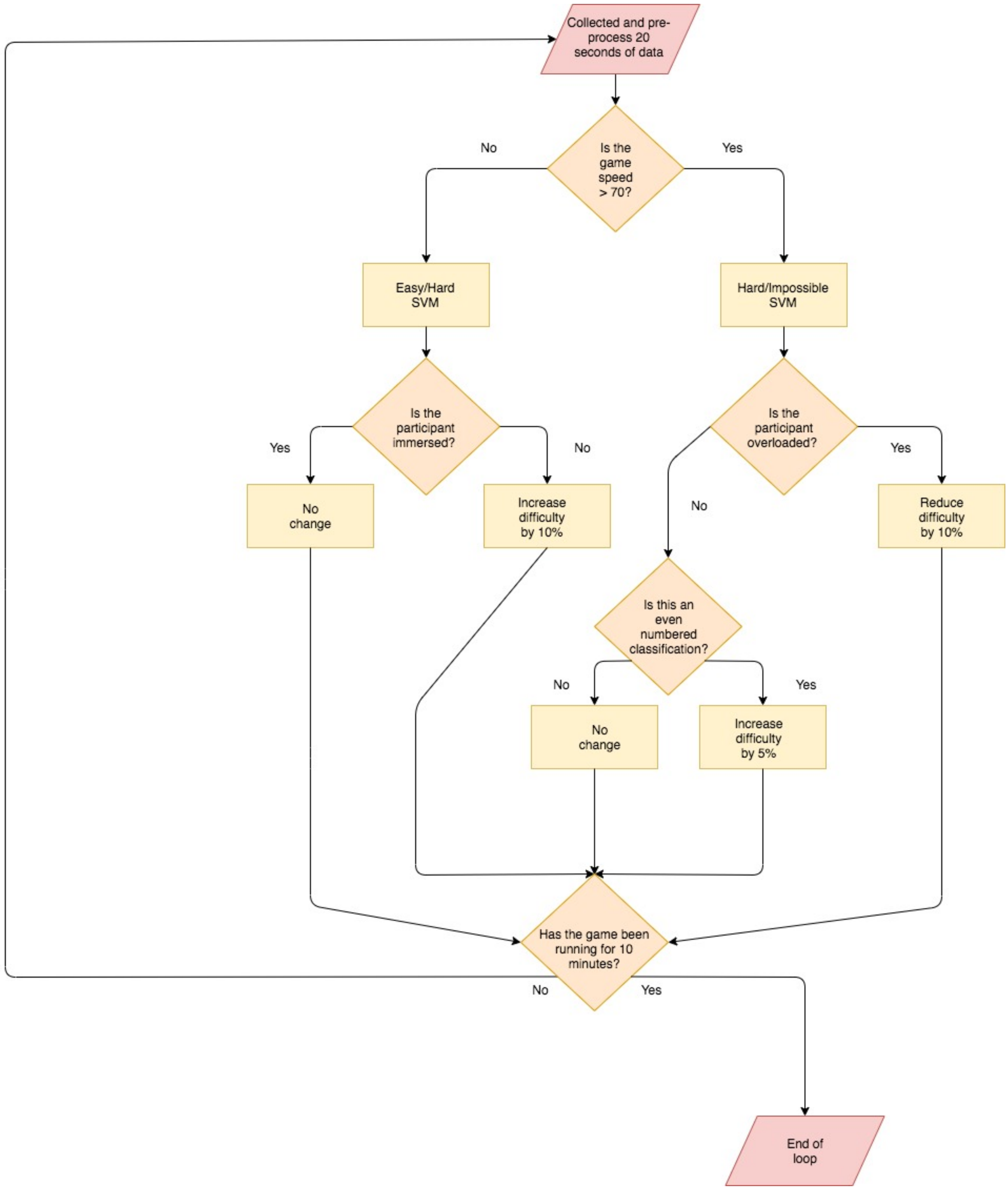


Figure 51 - Flow chart depicting the process carried out to determine required changes in game difficulty during the neuroadaptive condition

6.2.5 DDA via a random game

A random adaptive game was designed as the control for this experiment to ensure that indications of improved pain tolerance were not due to potential boredom that may stem from playing a fixed-level game that was either too easy, or too difficult. Using a random adaptive game allows us to determine whether potential changes in both pain tolerance and the reported level of pain were due only to the neuroadaptive element of the developed game. Pain perception reduction may occur due to the overall experience of playing a game, so the random game was developed to account for this potential effect.

During the random adaptive game, the same hardware was used as in the neuroadaptive game (Alienware laptop, MacBook Air, Sony PlayStation 3 Remote, Oxymon Mk II fNIRS kit.) The game would either remain at the same level or become harder or easier based on a random decision. To ensure that the two games (the neuroadaptive game and the random adaptive game) were comparable, the random adaptive game and the neuroadaptive game would trigger changes at the same frequency. This was to ensure that the two games produced comparable results, whilst still remaining different in their overall functionality.

During the random game, the game would start at the easiest level (speed 30) and remain at this speed until the first classification had been made. In keeping with the neuroadaptive game design, classifications were performed every 20 seconds. fNIRS results were collected from the participant and classified during the random game, although these results did not inform the random adaption. The purpose of collecting these results was to determine whether the random game was making correct or incorrect changes so that the effects of the random game could be more accurately compared to the results of the neuroadaptive game. When the game was in the Easy/Hard band (30-70), this adjustment would either be ‘No adjustment’, or ‘Increase difficulty.’ If the difficultly was increased, this increase was by 10%, in keeping with the neuroadaptive game. If the game was in the Hard/Impossible range, the adjustment would be ‘No adjustment’, ‘Increase difficulty’ or ‘Decrease difficulty.’ Decreases in difficulty were by 10%, but increases could be either by 5% or 10%. The percentage increase was chosen at random. During the random game, the decision of how the game should be adapted was made using the MATLAB function *rand* and was not based on results from either fNIRS or current game score. A flow chart depicting this process can be seen below in Figure 52.

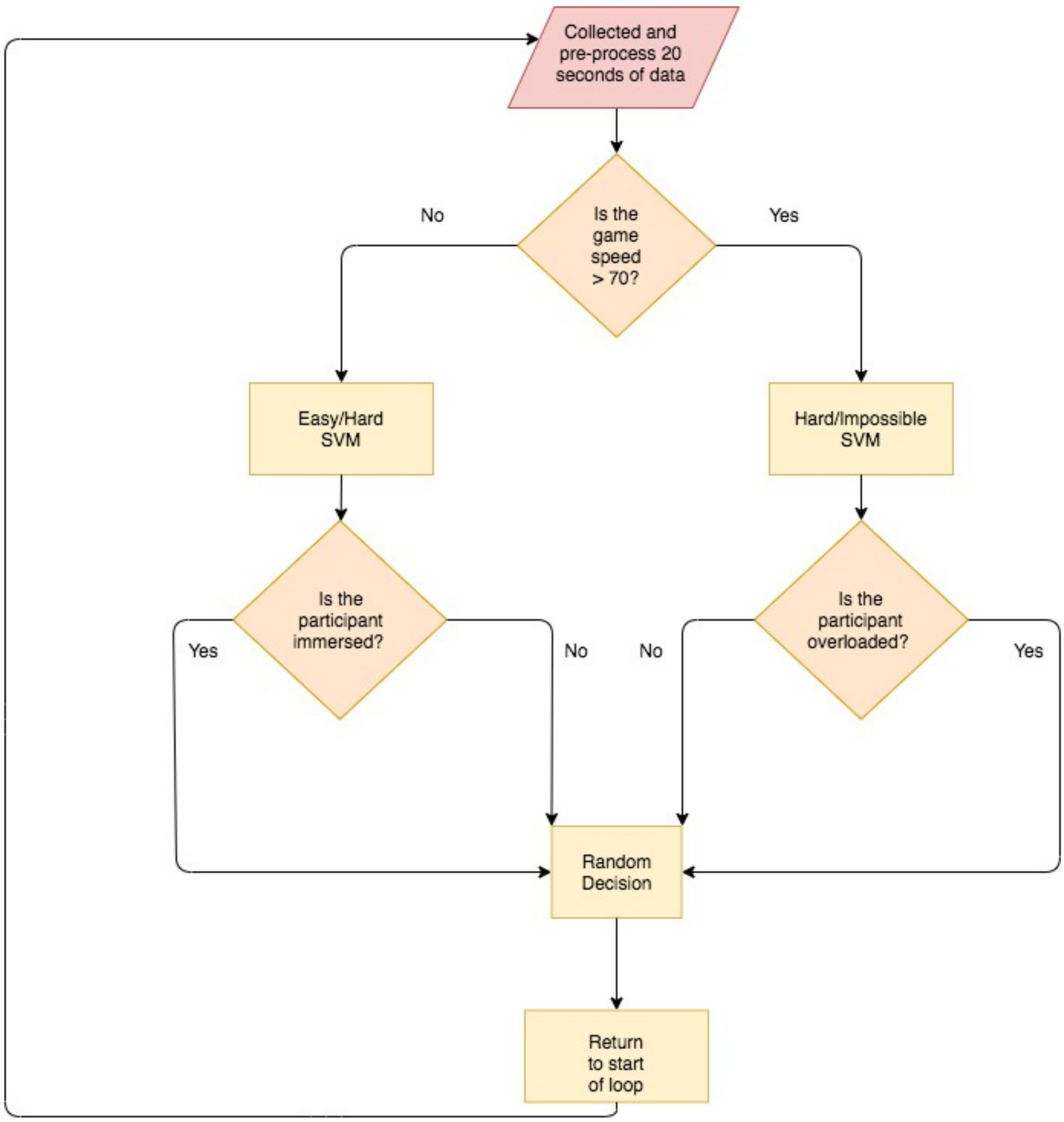


Figure 52 - Flow chart depicting the process carried out to determine required changes in game difficulty during the random condition

6.2.6 Materials

The Cold Pressor Test (CPT) was used to induce experimental pain. The water temperature for the CPT and protocol were identical to Chapters 2 and 3. The timings of participant Cold Pressor submersion were recorded to measure changes in pain tolerance. The Visual Analogue Scale (VAS) was used as a measure of pain perception. The NASA Task Load Index (TLX) (REF) was used as a measure of the level of subjective mental workload experienced by participants. The Motivation scale was used as a measure of intrinsic motivation. The VAS, TLX and Motivation scales were completed after each game condition. The subjective questionnaires used in this study are discussed in more detail in Chapter 3.

6.2.7 Procedure

Prior to the start of data collection for Study 3, a pilot test was carried out to ensure that the real-time protocol for data pre-processing and classification were functional. Before the experiment began, each participant completed a familiarization with both the game, and the CPT. This was to ensure that the data that was collected was an accurate representation of the participants pain tolerance and gaming ability and was not affected by being unfamiliar with the stimulus. Following familiarization, the participants completed a timed CPT, without a distraction, to measure their standard pain tolerance. This control was carried out to ensure that an accurate measurement of the difference in pain tolerance between the two game conditions could be established

The experiment began with a 30 second baseline, where the participant was asked to remain calm and prepare them self for the experiment. After this baseline period, one of the two games (random adaptive or neuroadaptive) was chosen using the MATLAB *rand* function, to ensure that neither the participant nor the experimenter would know which game was being played first. This information remained blind until the end of the experiment session. The first part of the experiment consisted of the participant playing the game for 5 minutes, without experiencing experimental pain. This five-minute adjustment period was established to ensure that the game had been adapted to the optimum level of difficulty before the CPT began. After the five-minute adjustment period, the participant was asked to place their foot into the Cold Pressor to begin the CPT. The participants foot submersion was timed from the point in which their foot first entered the water, to the point where the foot was removed. This time was recorded by the experimenter, and the participant continued to play the game until they had been playing for 10 minutes in total (5 minutes for game optimisation, and 5 minutes for playing at the optimal level of game difficulty). Regardless of how long the participant was able to keep their foot in the water, they would always play the game for a total of ten minutes.

Following the completion of the first condition, the participant was given the opportunity to dry their foot and warm it up, to ensure that any pain they may still be experiencing from the CPT was alleviated. The participant was then asked to complete the VAS to document their pain at its worst point. Following this, the participant was asked to complete the TLX questionnaire to establish their level of engagement with the game that they had just played.

The second part of the study consisted of the participant playing whichever version of the game that they had not already played, depending on which game was randomly assigned as the second condition. Again, the second game would last for a 10-minute period, 5 minutes for level optimisation and 5 minutes for the CPT. The participant always continued to play the game until they had been playing for 10 minutes, regardless of the length of their CP submersion. Again, following this 10-minute condition period, the participant was able to warm and dry their foot before completing the VAS and TLX.

Once the participant had completed the two game conditions and the questionnaires, they were debriefed, and the experiment concluded. The results of the study, including CPT times and level adaptations, which were made during each game, will be discussed further in the following section.

6.3 Results

The results of this study have been analysed with reference to measures of pain, subjective mental workload, and the adaptive behaviours of both the random adaptive and neuroadaptive game.

6.3.1 Pain

A one-factor repeated measures ANOVA illustrated that the main effect of the game condition was statistically significant [$F(2,17)=3.53$, $p=.05$, eta=0.29]. Post-hoc tests indicated that the effect of playing either game caused a statistically significant increase in pain tolerance compared to the baseline (no game) (Figure 53). However, there was no significant difference in CPT times between either form of the game.

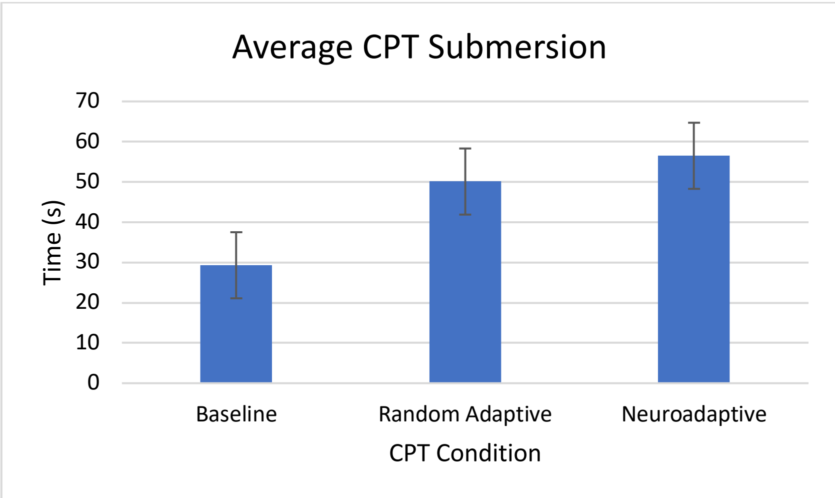


Figure 53 - Bar graph displaying the average and standard deviation for CPT submersion over three conditions: Control, Random and Neuroadaptive

Table 13 presents the measurements (cm) of the VAS that each participant completed after each CP immersion during a gameplay condition. A one-factor repeated measures ANOVA illustrated that effect of the game condition fell outside of statistical significance ($F = 0.18$, Hypothesis DF = 1, Error DF = 19, $p = .894$, Eta .001.).

Table 13 - Results of the VAS measurement (cm)

	Adaptive	Random
Average	**13.34**	**13.45**
Standard Deviation	**4.17**	**4.39**

The results indicate that there was no significant difference in reported levels of pain between the two game conditions.

6.3.2 Subjective Mental Workload & Motivation

Table 14 presents the average results of the TLX and Motivation questionnaires. A one-factor repeated measures ANOVA of the TLX results illustrated that effect of the game condition fell outside of statistical significance (F = 0.13, Hypothesis DF = 1, Error DF = 19, p = .911, Eta .001). Therefore, there was no significant difference between the two game conditions with respect to subjective mental workload.

Table 14 - Results of the TLX and Motivation questionnaires

	TLX		Motivation	
	Random	Adaptive	Random	Adaptive
Average	**4.53**	**4.50**	**39.10**	**40.65**
Standard Deviation	**1.00**	**1.65**	**4.24**	**4.71**

In terms of the motivation scores, the level of subjective motivation was significantly higher in the Neuroadaptive game as opposed to the random adaptive game. A one-factor repeated measures ANOVA illustrated that effect of the game condition was statistically significant in this case (F = 5.535, Hypothesis DF = 1, Error DF = 19, p = .03, Eta .226.)

These results indicate that, although the participants were motivated more in the neuroadaptive game than the random adaptive game, there was not enough overall differentiation between the two gaming conditions to elicit a significant change in the results of the TLX.

6.3.3 Behaviour of the System

The difficulty of the game was adjusted upwards and downwards in real-time by manipulating the speed of the game. The speed of the game varied between a minimum of 30 and a maximum of 160. Both the Neuroadaptive and the random-adaptive game could adjust the speed by 10% when the speed of the game was lower than 70. When the speed of the game was greater than 70, speed could be adjusted upwards or downwards by either 10% or 5%. The first Figures 54-56 display the frequency and direction of difficulty change that occurred to the game at each level of difficulty. The graphs indicate the comparison between the neuroadaptive and random adaptive game. Figure *54* shows the frequency of all upward increases across the full range of game speeds for all participants, using both neuroadaptive and random adaptive games.

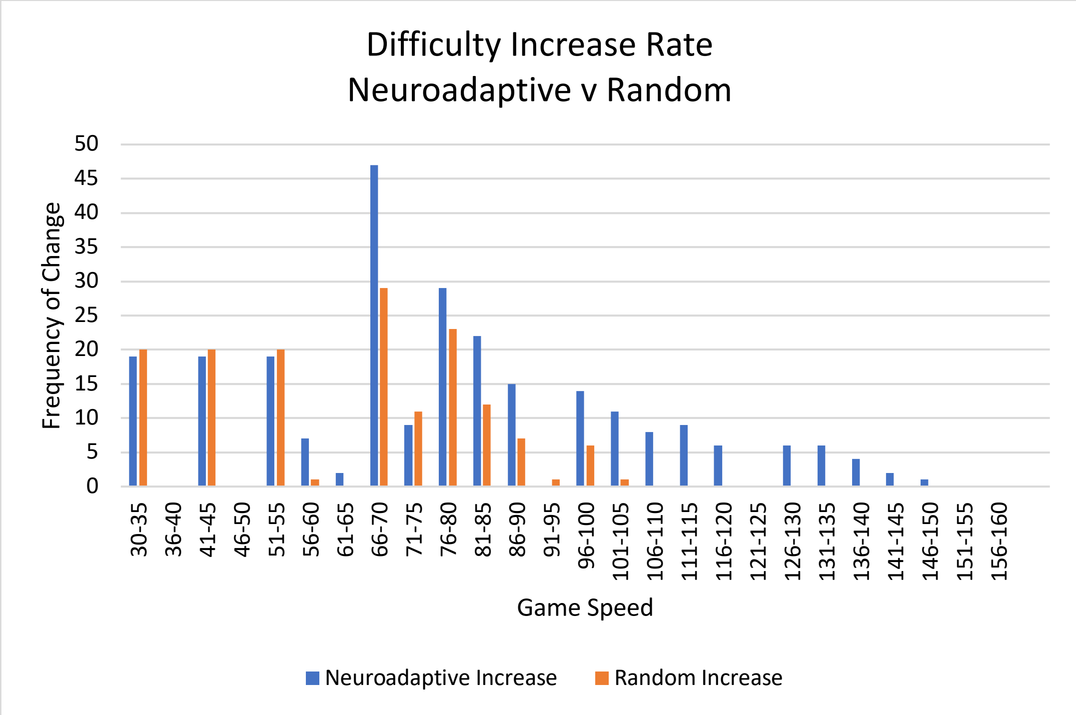


Figure 54 - Frequency of adjustments to increase game difficutly at varying speed for neuroadaptive and random game

Figure 54 indicates that the neuroadaptive algorithm increased game difficulty more frequently between the speeds 66-77 compared to the random adaptive game. This figure also displays that the neuroadaptive game continued to increase the level of difficulty at levels above 101, whereas the random algorithm did not. The frequency of downward adjustments when game speed was greater than 70 is illustrated in Figure 55.

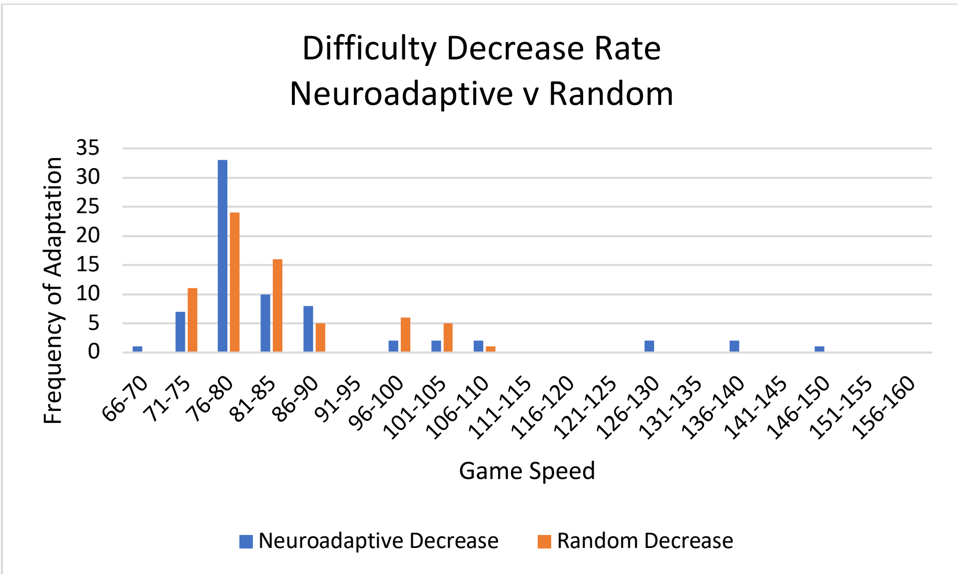


Figure 55 - Difficulty decrease rate at varying speed for neuroadaptive and random game

As displayed by Figure 55, the difficultly decrease rate was largely similar between both the neuroadaptive and the random adaptive game. However, a similar pattern between the difficulty increase and decrease can be found in terms of the neuroadaptive game continuing to decrease past a level of 126, whereas difficulty ceases to decrease at level 110 in the random game.

Frequency of periods where no adjustment was made is shown in Figure 56. When game speed was greater than 70, a classification of immersion would not trigger an adjustment. If the participant was not immersed, the game speed was increased, and no subsequent adjustment was made for the next potential adjustment period. When the game speed was less than 70, either upward or downward adjustment automatically lead to no adjustment in the subsequent period.

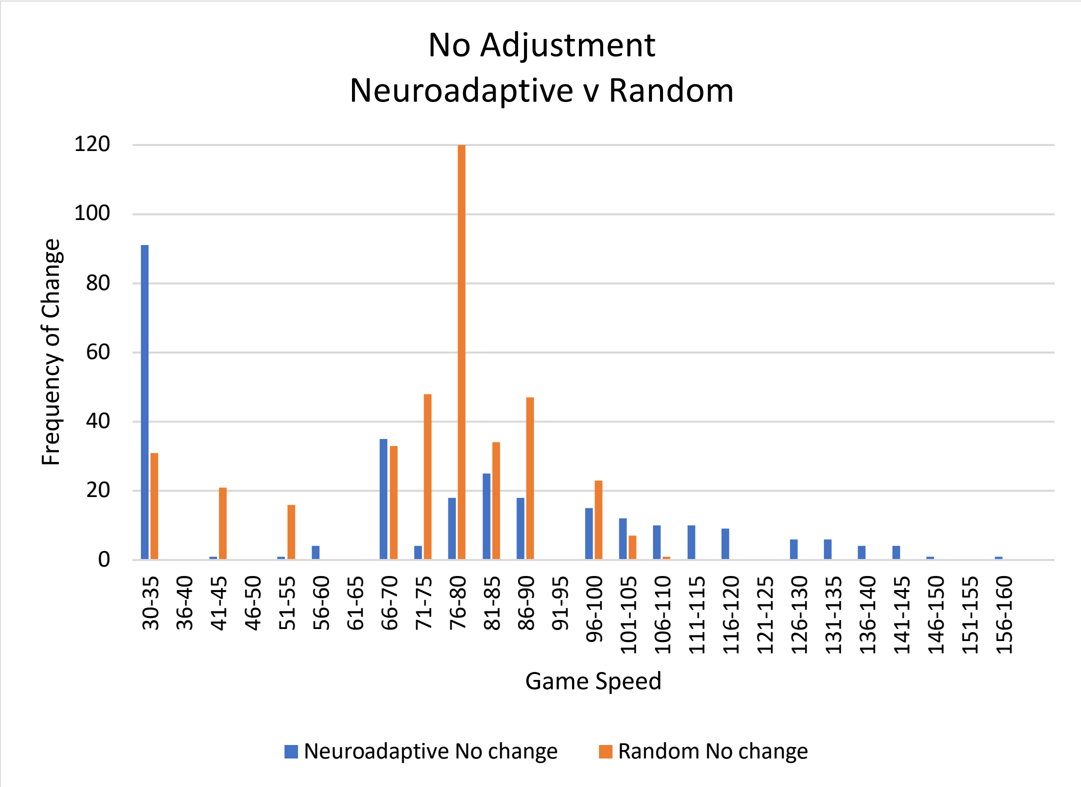


Figure 56 - Difficutly remain rate at varying speed for neuroadaptive and random game

Figure 56 depicts the frequency of level maintenance between the random adaptive and the neuroadaptive game. It is clear to see that the neuroadaptive game sustained difficultly at the lowest level far more frequently than the random game did. In all cases, apart from level 30-35 and level 66-70, the random game far more likely to apply the 'no change' option than the neuroadaptive game was.

Figures 47 – 49 present the number of adaptations, which were made at each level of game difficulty in both the random and the adaptive condition. These figures represent a culmination of the adaptations that occurred for each participant. These figures indicate that adaptation continued to occur up to the level of 160 during the neuroadaptive game, whereas adaption ceased after the level of 110 for the random game. These results indicate that there were noticeable differences between the functionality of the neuroadaptive game and the random game. The individual changes that occurred per-level for each participant are displayed on the following page in Figure 57.

Figure 57 displays the level of game difficulty for each participant during the random and the neuroadaptive game. The vertical blue line on the graph indicates the point whereby 15 adaptions have been made. After five minutes (equal to 15 adaptations), the participant was asked to place their foot into the CPT. At this point, the game should have been at the optimal level for distraction from pain. This figure shows that, although some participants were playing the game at drastically different levels (particularly participants 14 and 20), there was a lot of similarity between the levels of game for many of the participants, particularly at the point where the CPT began. Overall, Figure 57 indicates that the level of game difficulty is often similar regardless of which adaptive game the participant was playing. A one factor repeated measures ANOVA indicated that there was no statistical significance found between the levels of difficulty at the start of the CPT [$F(3,14)= 0.392$, $p = .539$] regardless of which adaptive game the participant was playing.

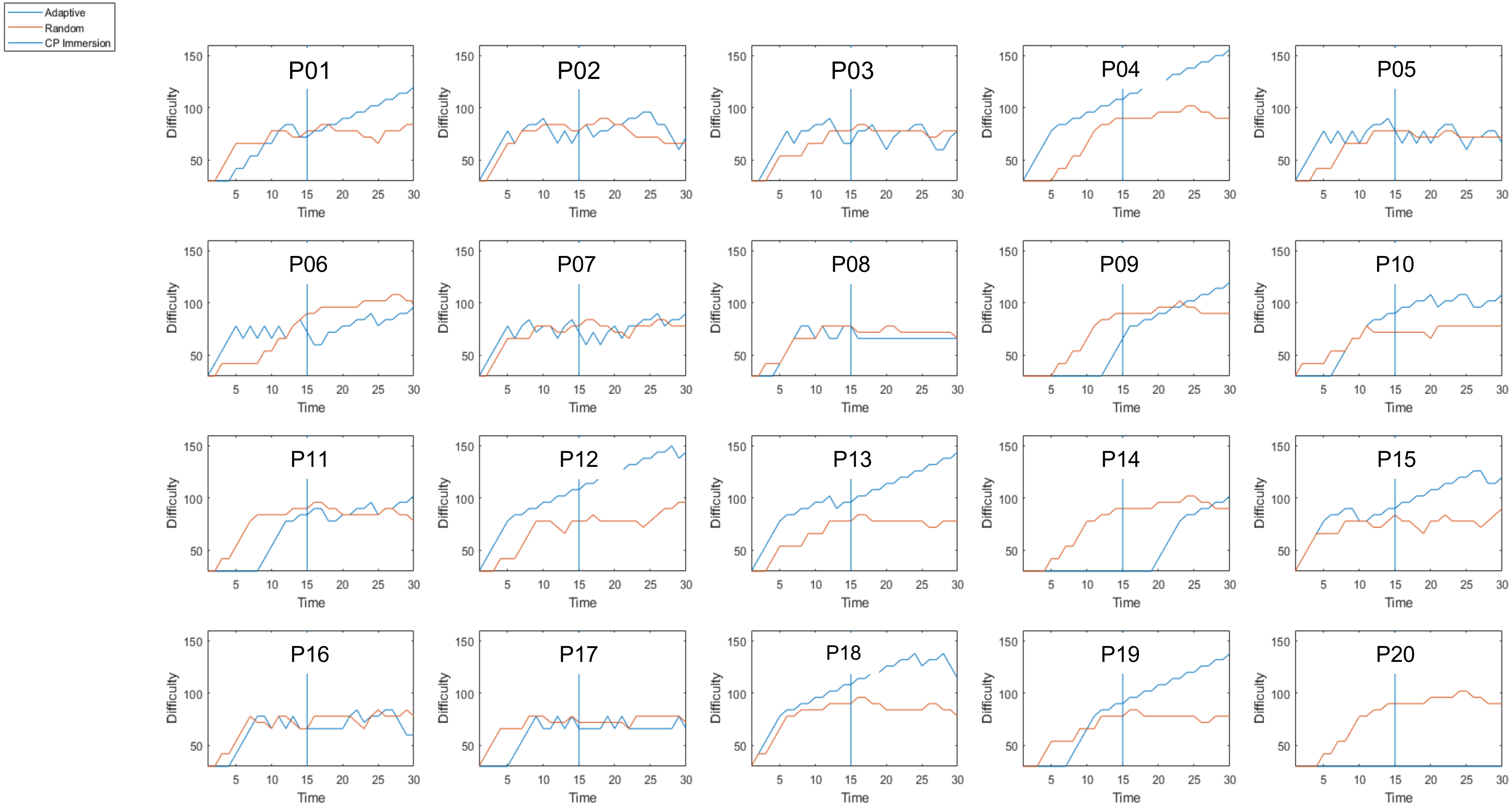


Figure 57 - Level of difficutly throughout gameplay for both neuroadaptive and random game. Scale (1-30) represents the amount of neuroadaptive changes which were made

6.4 Discussion

The purpose of this study was to assess the viability of the neuroadaptive game as a non-pharmaceutical analgesic. We expected the results of this study to indicate that a neuroadaptive game would reduce subjective pain and increase pain tolerance to a significant degree, in comparison to a random adaptive game. It was expected that a neuroadaptive game would maximise attention and therefore produce these results. However, we found no evidence to support any of those claims in the results, as both systems (neuroadaptive and random adaptive) produced similar outcomes. The results of the CPT and VAS indicate that there was no statistical significance found between the random or neuroadaptive game. The CPT and VAS results provides respective indexes of pain tolerance and subjective pain. The results of the TLX questionnaire, which measures the subjective workload of the participant, were also statistically insignificant between the two games. Therefore, there was no evidence that either game was experienced as more demanding or effortful than the other one, i.e. the neuroadaptive rationale did not significantly increase mental workload above the system that behaved randomly. The only result that achieved a level of statistical significance ($p = .03$) was the result of the motivation questionnaire. When considered together, the TLX and Motivation results indicated that the overall gameplay experience was not perceived as different by the participants, but participants did experience increased subjective motivation whilst playing the neuroadaptive game. It could be argued that the reason for this statistical significance was that, although the two games were not particularly different in terms of the level of demand, participants were more likely to be more motivated to play a game where the difficultly level was tailored to their personal ability via dynamic neurophysiological markers of attention.

A visual inspection of difficulty levels (Figure 53) indicated that both random adaptive and the neuroadaptive game were very similar throughout gameplay for most participants and no significant difference was found between the levels of difficulty of the two adaptive games at the point of CP submersion. Therefore, it is unsurprising that no statistically significant effect was found on pain tolerance between either of the games because game demand is a primary driver of pain tolerance (Fairclough et al (Under Review)). Although the neuroadaptive game did continue to alter the difficulty of the game beyond a level of 105 (where the random adaptive game did not,) the frequency of adaption was largely similar across both games at each individual level. There was a trend in the 'No Change' graph (Figure 57) wherein the neuroadaptive game had a far larger amount of 'No Change' decisions at the lower level (30) than the random adaptive game.

Figure 57 displays the level of difficulty that the participant was playing at the point where the CPT began. In a portion of these results, it is clear to see that the level of difficulty was not as varied as we had hoped, especially at the point of foot submersion. We hypothesise that the lack of statistical significance found in the objective and subjective results could be largely due to the fact that the two versions of the game were too similar to each other to provide an accurate comparison. Game speed or demand was almost identical for both the neuroadaptive and random adaptive games at the point of CP submersion. For this reason, we would not expect to see any differentiation between the two systems with respect to pain tolerance. The absence of statistical significance for CPT, subjective pain and workload can be interpreted in one of three ways, either: (a) the neuroadaptive game was ineffective, (b) the random adaptive game was more effective than anticipated or (c) the neuroadaptive game was ineffective and the random adaptive game was more effective than anticipated. In any of these cases, the behaviour of the systems indicated that both experimental conditions were not clearly differentiated from one another; a factor that makes the experimental results difficult to interpret. The absence of any differentiation between the two systems, with respect to the adaptation of game speed, points to a broader methodological issue concerning the evaluation of technological systems that adapt in real time.

In order to effectively evaluate the efficacy of a neuroadaptive game, there must be an alternate gaming condition to provide a point of comparison. As can be observed by the results of the CPT submersion times (Figure 53), playing any type of game is effective in increasing pain tolerance when tolerance is measured between game and no game conditions. The issue here is determining how to effectively compare a neuroadaptive game to a 'standard' game and what constitutes a reasonable point of comparison for a neuroadaptive game. It is already understood that a game that is either too easy or too difficult will not engage the attention of the player. We also understand that attentional engagement represents an interaction between the skills/motivation of the individual and the objective level of task demand, and therefore, the optimal level of demand must be personalised to the individual. For this reason, it would be unsuitable to compare a neuroadaptive game to a fixed level game, as individual differences in gaming ability and preference could skew the results in favour of the neuroadaptive game, due to boredom or cognitive overload, but not because a neuroadaptive game was actually more effective than a non-neuroadaptive game. The variation inherent to the neuroadaptive game provides it with an advantage compared to fixed and unchanging levels of game demand. The same issue could arise in a truly random game, wherein the game level can change across a wide spectrum with no consideration for the previous level of game difficultly (i.e. in a non-linear fashion). Generally speaking, adaption during gameplay through DDA is usually represented in a linear fashion, to allow the player to gradually adjust to increases or decreases in difficulty (Zook 2012). Therefore, unpredictable changes in difficulty are not only uncommon in most games, but could also reduce the overall experience of immersion, as the participant would be unable to adjust to each level of difficulty before the difficulty was changed again. Changes in game demand that are random and unpredictable can reduce attentional engagement of the player by their probabilistic nature, but equally, some players may be entertained or engaged by this kind of randomisation. Designing a game such as this could skew the results in favour of the neuroadaptive game, not truly because neuroadaptation is effective, but rather because it is more effective than a completely random game. For the reasons outlined, it is fundamental that the neuroadaptive game is compared to another adaptive game, but this game must not have been designed with the sole intention of being unpredictable and difficult to play.

Creating an adaptive game that could be used as a reasonable comparison to the neuroadaptive game would mean that the two games had to be clearly differentiated from each other. The game that was created for this study was a random adaptive game, but the game still followed a linear and controlled pattern of difficulty increase/decrease to ensure that the game was not entirely unpredictable. The schedule of adjustment for the random game was not random as it was controlled by the same

SVM as the neuroadaptive game. The only difference between the two games was that the random version made a random adjustment, whereas the neuroadaptive game made an informed adjustment based on the results of the SVM classification. Although the ‘random’ game was adapted randomly, there was still a 50% chance in the Easy/Hard band, and a 33.3% chance in the Hard/Impossible band, that the correct adaptation would occur each time an adaptation was made. In hindsight, using a random game where the initiation of game difficulty adjustment was random (based on a fixed probability) with the second level of randomisation (whether difficulty adjusted upwards or downwards) used in the current study may have provided a stronger comparison. The results of this study indicate that, although the two game conditions followed different adaption patterns, they were very similar to each other, especially at critical points, such as the beginning of the CPT. The results of this study indicated that a different route for creating a comparison condition would be required to efficiently test the effectiveness of a neuroadaptive game. A second alternative would be to design a yoked control, wherein adaptations are made to the game for Person A, based on previously collected data from Person B. This would mean that the two games would follow the same adaptive logic, but only one game would be neuroadaptive to the specific individual playing the game at that time as the yoked condition would represent required changes based on live neurophysiology, but not the activity of the person who was currently playing the game. On reflection, creating a game that adapted in this way may have been more effective than the random game that was created in this study. One alternate route that could have been followed would be to retain the adaptive logic that was used in the neuroadaptive system but utilise a different category of measurement, such as game score or heart rate. This control would allow for an adaptive comparison that still followed the typical rules of a DDA system, but perhaps yielded less accurate classification results than the full neuroadaptive game, and therefore the level of difficulty would be less suited to the participant. However, if the classification relied only on game score, then it may have taken longer to classify the immersion of the participant in the Easy/Hard band, which would have, in turn, increased the amount of time it took for the participant to reach the desired level of difficulty. Had the game been created in this way, we may have achieved more significant results, as it is reasonable to expect that the control game wouldn’t have been as effective in retaining the participants attention. Furthermore, the time of the adaptive window (20 seconds) over a 10-minute game only allowed for 30 decision points in the whole game and therefore only 15 changes prior to the CPT. It is possible that 30 adaptations (or indeed 15 pre-CPT) were not sufficient for finding the optimal level of difficulty. There are several examples from Figure 57 (P01, P04, P10, P12, P15 and P18) where differences between the neuroadaptive and random adaptive game only begin to occur after 15 adaptions, at which point the CPT has already began. Upon analysing the results of Study 3, it has become clear that the two games were too similar to provide a contrasting comparison. However, the results of the Motivation questionnaire do indicate that participants were aware (at least in part) of the differences between the two games. We theorise that, had a different route been followed for the creation of the adaptive game (for example, the levels changing based on game score or heart rate,) we may have been able to achieve stronger results.

Moving forward, if this study were to be replicated, we would suggest that further testing is carried out during the creation of the random game. If statistical significance could be achieved between the two game conditions (random and neuroadaptive), we expect that the results indicting the perception of pain may also be more significant. It would be interesting to compare different versions of a non-neuroadaptive game (for example, a game that relies on heart rate data, compared to a game that remains at a fixed level, compared to a game with a yoked control) to find which of these games provides the most statistically significant results compared to a neuroadaptive game. Once the ideal control condition had been found, repeating this study using the original neuroadaptive game and the new control condition may provide a clearer indication of the overall effects of a neuroadaptive game. Although the results attained in this study did not prove our hypothesis, we still consider this experiment to be a useful evaluation exercise. Completing this study has enabled us to determine that further development of a random control condition is required to fully assess the effectiveness of the developed neuroadaptive game. If further studies were carried out with an adapted version of the random game, and statistical significance was still not achieved, it would indicate that further development of the neuroadaptive game (or of the style of game entirely) is required to fully examine the hypothesis.

As mentioned in Chapter 5, we made the decision to use SI classification in order to streamline our data collection process – using SI classification allowed the neuroadaptive game to work as a standalone system without any personalised training required. Although we do think this is the best option for our purpose (when considering a potential move to a clinical setting), it would certainly be worthwhile to create and test a SD system to determine how differences in classification may have been affected. It is possible that relying on SD classification rather than SI classification may have improved our results overall, however we do feel that the overarching issue remains the control condition (random game) and SD classification should only be considered if changes to the control condition do not yield satisfactory results.

In summary, this study found no statistical significance in pain tolerance nor pain perception between a random adaptive game and a neuroadaptive game. Mental workload, as measured by the TLX, was also statistically insignificant between the two conditions. The only recorded measure where statistical significance was found was in self-reported motivation scores. Upon analysing the results, especially when considering that both games were effective in increasing pain tolerance compared to a baseline no-game condition, there is a strong indication that the flaws of the study design may be the reason for the lack of statistical significance.

Chapter 7 - Discussion

7.1 Introduction

The studies that were carried out during our research were not able to prove that there can be a statistically significant improvement in pain tolerance when a neuroadaptive game is used. However, we did uncover some results that indicate why this may have happened, and what further work could be done to thoroughly evaluate the hypothesis.

7. 2 Main Findings

The hypothesis of this cumulative research project was as such: *does increased game demand increase pain tolerance*? This hypothesis formed the backbone of the research, and Studies 1 and 2 (Chapters 3 and 4) indicated that, although the results were not statistically significant, there is an observable increase in the trend of Cold Pressor Test data, which indicates that an improvement in pain tolerance can be achieved as the demand of a distraction task increases. Although the results of Studies 1 and 2 were not statistically significant, the trend of the data indicated that it was worthwhile to carry out a further study and determine whether neuroadaptive game difficulty may provide a statistically significant effect on pain tolerance when compared to a non-neuroadaptive game. The data collected during Studies 1 and 2 were used to inform the creation of the neuroadaptive game, which was used in Study 3. Studies 1 (Chapter 3) and 2 (Chapter 4) were performed to determine the optimal feature set that could be used for real-time classification of game difficulty. Both studies revealed that features derived from fNIRS data were consistently superior to alternative data types (e.g. ECG-based features, game scores) in relation to classification accuracy. In Study 1, the best SVM classification for game demand using HR data was 55.4% compared to 86.4% for fNIRS. In Study 2, the use of Game Score data for classification of game demand did provide more accurate classification results in comparison to the HR data. For the Easy vs Impossible classification, game score data was able to classify at a similar level of accuracy as fNIRS data. However, fNIRS classification (86.4% classification accuracy) was superior to Game score classification (63% classification accuracy) for the Easy v Hard condition. We speculate that the poorer classification rate achieved by the game score data during the Easy v Hard condition was due to the distance between the conditions, i.e., there is a smaller distance between Easy and Hard conditions, which could lead to a reduced difference in game score variation. However, there is a larger difference in gameplay between the Easy and Impossible conditions, which would lead to a larger score variation, and more successfully classification. These results indicate that fNIRS data is more sensitive to small changes in game difficulty than the game score data itself.

Previous research into the use of gaming systems for the distraction from pain (to increase pain tolerance or reduce pain perception) largely focuses on the comparison of a gaming technique (active) to an alternate passive technique. The research discussed in Chapter 1 indicates that an active distraction technique will always provide a more effective distraction from pain than a passive distraction technique. There is little research that compares the differences between alternate active distraction techniques. Our final study (Study 3) compares two types of active distraction – a neuroadaptive distraction and a random distraction. We believe that the use of two active distractions could be partially responsible for the lack of statistical significance within our results. It would be useful for our work if research was carried out which focuses on the differences between alternate active distraction techniques. This would provide further results which may indicate how effective a neuroadaptive system can be. It is important to note that our neuroadaptive system was more effective than no distraction – which is a positive result and indicates that the distraction was sufficient. The lack of research between alternative active distractions does make it difficult to draw conclusions as to weather a neuroadaptive system is more effective than a standard game – as there is little research which compares different active distractions with a view to distracting from pain.

It is important to note that a large proportion of the research into gaming for pain relief uses genuine clinical settings to assess the response to pain. The biggest difference between pain induction techniques such as the CPT and genuine clinical pain is that the participant in CPT studies is in control of their pain. Participants therefore do not have to endure pain for longer than they feel able to. In clinical settings, procedures have to be carried out to completion regardless of the level of pain of the participant. Every effort is taken to reduce pain and discomfort in clinical settings, but patients largely do not have the ability to stop the painful experience. It is possible that our results have been affected by the participants ability to end the pain experience whenever they wish. It may be prudent to test our current system (and its control counterpart) in a clinical setting where pain cannot be controlled by the participant/patient, to determine whether our neuroadaptive system may already be affective in clinical settings.

Although fNIRS data is less practical to collect than either HR or Game Score, the benefits of collecting and classifying fNIRS are considerable. Figure *58* below indicates the combined results from Studies 1 and 2. The nature of Space Ribbon, used only in Study 1, means that no game score data could be collected in this case – and the results of Study 1 indicated that HR results were not suitable for classification. For this reason, HR is provided for Study 1 only, and Game Score is provided for Study 2 only. fNIRS data is provided for both studies 1 and 2.

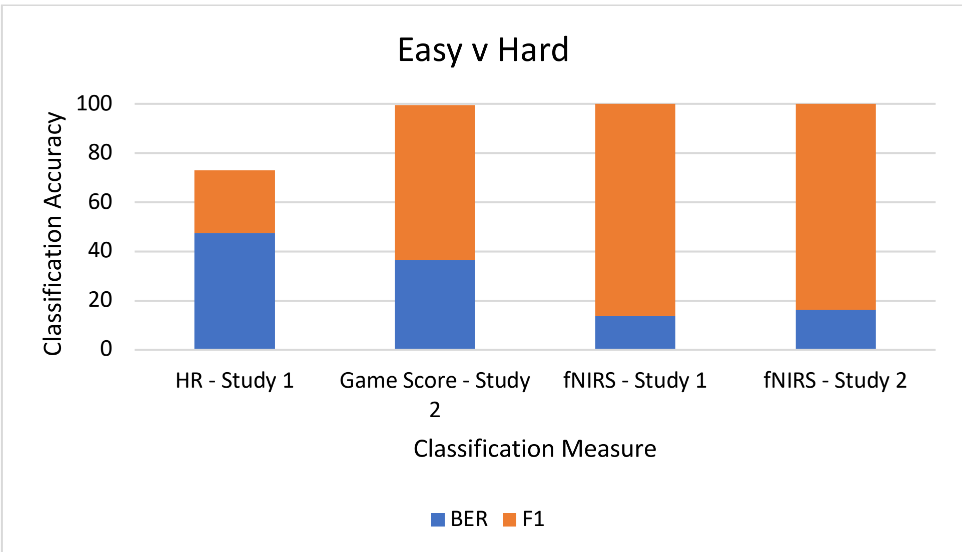


Figure 58 - Combined classification results across Studies 1 and 2

The results in Figure 58 illustrate the reduction in BER and increase in F1 score that can be found when fNIRS data is used as compared to HR (Study 1) and Game Score (Study 2). This figure indicates that using a combination of fNIRS data and game score is the most accurate dataset for classification overall. There are various reasons as to why fNIRS data can provide a more accurate classification than game score data. Firstly, fNIRS is sensitive to changes in attention at cortical sites, whereas the game score data only provides data relating to in-game achievements. Although the achievements of the participant may be similar across two difficulty levels, it is possible that a participant is exerting more attention during a more difficult game level, in order to achieve the same score as they previously achieved in an earlier level. These changes, although not observable in game score data, are observable in fNIRS data. Secondly, fNIRS is collected from a variety of channels, which means that it produces a discrete and cumulative measure of connectivity across multiple brain areas. fNIRS can indicate how the relationship between areas of the brain are affected by changes in game demand, and therefore provide a clearer indication of how game demand is affecting a participant's attention. Thirdly, fNIRS was collected at a rate of 10Hz, meaning that 10 data points were collected every second. In contrast, the game score data was only collected every 2 seconds. This meant that fNIRS provided a more sensitive indication of attention than what could be inferred from the game score. The benefits of using fNIRS is that changes in brain activity are more discrete and frequent than changes to the game score, which allows for a more accurate classification.

A further benefit of using fNIRS is that a large variety of features can be created from the filtered signal. Using feature selection, we can determine which created features provide the most relevant information for classification. The results gathered over the two studies (Studies 1 and 2) indicate that neuronal classification is the most sensitive method of adaptive game creation, regardless of the null results that were found relating to the increase in pain tolerance. Although we were unable to prove our main hypothesis, classification results (across Studies 1 and 2) indicate that classification of fNIRS data was accurate both in and of itself, compared to autonomic data, such as HR. Although fNIRS is a more expensive and less practical data collection technique in comparison to other psychophysiological methods (such as HR), HR alone is not suitable for classification of attentional state (especially in the case of video game studies). Video games are understood to increase the heart rate of a participant (Gwinup, Haw, & Elias, 1983) which means that it is difficult to gather an accurate baseline recording of HR. This is because the baseline conditions will be generally lower than any video game condition, whether the game is providing an engaging experience or not. A generalised increase in HR activity means that it is difficult to determine whether changes in heart rate are due to the difficulty of the game or just the overall gaming experience. Our HR classification results reflect this theory, as the HR classification was the poorest classification result attained. fNIRS is capable of measuring attentional regulation, whereas HR measurements provide no indication as to the reasons for potential fluctuations. Although we may be able to observe a change in HR, we cannot prove the cause of said changes. However, using fNIRS enables us to target areas of the brain that specifically relate to attentional regulation, meaning that we can verify the cause of changes in the fNIRS data.

During feature selection, there were more instances of connectivity measures that were relevant for classification than any other type of feature. This indicates that connectivity measures are inherently sensitive to changes in game demand. Connectivity features appeared more relevant for classification in both Studies 1 and 2, which indicates that these features are more indicative of attentional level, regardless of the fNIRS montage that is used. The fact that these features were more indicative of attentional state than discrete features indicates that, although discrete channel data can fluctuate, these fluctuations are only relevant if they coincide with fluctuations at other channels. For example, if we had observed an increase in HbO at one site, we could assume that this indicated higher activation in that brain area. However, if this increase in HbO at one site is accompanied by a simultaneous decrease of HbO at another site, there are implications that these data indicate changes in neuronal behaviour, which are assumed to relate more specifically to changes in attentional control. This theory is supported by other fNIRS research, which found that changes in the demand of a working memory task (1-back, 2-back or 3-back) showed an increase in front-parietal connections, alongside a simultaneous decrease in inter-hemispheric frontal connections, as task load increased (Fishburn, Norr, Medvedev, & Vaidya, 2014). Research indicates that the strength of connectivity measures can be improved via the use of the CBSI algorithm. One study (CBSI Ref) found that using CBSI increased the efficacy of Prefrontal cortex (PFC) network connectivity in response to task demand. In comparison to bandpass filtered fNIRS data, further treatment using the CBSI algorithm reduce component unrelated neural activity, providing stronger PFC connections during a task (Racz, Mukli, Nagy, & Eke, 2017). It has been proven that connectivity occurs between brain areas during specific tasks. For example, Corbetta et al. (Corbetta et al., 2008) discussed the connections between areas of the brain that occur during task switching. This is true also of speech recognition. For instance, Liu et al. (Y. Liu et al., 2017) conducted a study that required participants to listen to a story, both in a language that they could understand, and a language that they could not. This study found that connectivity between areas of the brain relating to linguistic recognition was only significant when the participant was being spoken to in a language that they could understand. This further verifies the ability of connectivity measures to infer specific psychophysiological responses within the brain. This research indicates why connectivity measures were selected by feature selection techniques, and why they provided superior classification rates. This is largely because the correlation response is more likely to infer the correct game difficulty than the response that occurs from a single channel. Not only have we successfully classified fNIRS connectivity data in two independent studies, but the literature also suggests that connectivity measures can be used to quantify attentional state (Harrivel, Weissman, Noll, & Peltier, 2013).

With respect to machine learning approaches, we used a restricted method (LDA) and a more flexible approach (SVM). LDA classifications attempt to find a linear combination of variables that can allow for discrimination between two classes. Linear implies the linear separation between two groups of variables, and the group with the most representation along this linear separation represents the classification that has been made. Whereas LDA relies on the global dataset, a SVM classification is performed on a local dataset that lies closer to the group separation boundary (Xiong & Cherkassky, 2005). SVMs are particularly useful in cases where data are not evenly distributed or do not follow a linear pattern. Instead of relying on a linear separation being found between two groups, SVMs rely on the highest level of similarity between groups, allowing data of a similar kind to be more accurately separated and classified (Auria & Moro, 2011). This is especially relevant when considering fNIRS data, which can contain both linear and quadratic patterns. However, issues that can occur when machine learning techniques are used is that it can become difficult to separate data when there are more data points than there are dimensions ($N < p$).

We chose to use subject independent (SI) classification in Study 3 (the real time experiment), and therefore both Studies 1 and 2 also used subject independent classification techniques. There are many benefits of using SI techniques (as opposed to subject dependant (SD) techniques). The primary benefit specifically for our research is that, with SI classification, there is no individual training required for participants, which means that the neuroadaptive system can be used instantly for each new participant. This saves time and provides a more robust, user friendly system. SI classification also enables new data to be added to the classification after collection, which enables a larger and richer dataset to be used and improved upon each time the system is used, if desired. There are, however, drawbacks to such a system, which will be discussed in the following section.

7.3 Limitations

Throughout this research project, we did not find statistically significant evidence to indicate that pain could be reduced via game difficulty manipulation. However, we are confident (due both to previous research and the classification results that were attained) that this null hypothesis was not due to the use of fNIRS data as the control adapter for a neuroadaptive game. There were a number of limitations inherent within the experiments carried out during our research that should be discussed.

One issue that was present throughout all three studies is the use of the Cold Pressor Test (CPT) as a method of pain induction. The use of the CPT produced high variability within our results. During Studies 1 and 2, each gameplay session only lasted for three minutes. This short gameplay time meant that the temperature of the CPT had to be low, to ensure that there would be enough time for the trial to elicit a pain response. However, using a colder temperature meant that some participants found the water immediately painful and removed their foot after less than twenty seconds. Another issue, relating to the use of the CPT, was the point in which the participant placed their foot into the water. In Studies 1 and 2, participants were instructed to place their foot into the CPT immediately as the game condition began, which may have meant that the participant was not immersed in the game before the induction of pain, and therefore .the induction of pain may have prevented the participant from experiencing immersion. Alternately, although the CPT immersion began partway through the game condition in Study 3, the participant was instructed by the experimenter to place their foot into the water at this point, which may have meant that the participant was distracted from the game condition. Distraction from the game condition could disrupt the potential experience of immersion. Overall, although participants were asked to retain their foot in the water until the pain became intolerable, participant's personal perceptions of intolerable pain are likely to be widely variable, which may have meant that participants were not categorising their pain in the same ways. However, inducing pain via alternate methods to the CPT may have provided a more reliable experience. For example, a thermal pain protocol could have been used. A thermal pain protocol allows for a heat stimulation threshold to be calibrated for each participant, which means that at no point will a participant experience pain above their predefined and self-selected threshold. One benefit of using thermal pain, as opposed to the CPT, is that, because we can ensure participants will never experience a painful experience that they cannot withstand, participants do not require the ability to remove themselves from the painful stimuli during the trial. Using thermal pain would mean that participants could all experience pain for the same length of time during the experimental protocol, which may reduce variation in the VAS results, and also negate issues relating to the participant having control over when their foot entered and exited the CPT. A variety of studies have been carried out using a thermal pain protocol relating to changes in pain perception both for cognitive tasks (Rémy, Frankenstein, Mincic, Tomanek, & Stroman, 2003) and gaming experience (Czub & Piskorz, 2014)

A second limitation of the methodology was the nature of the game conditions themselves. We identified three (study 2) or four (study 1) levels of game difficulty and categorised these levels between Easy and Impossible. We chose to use a racing game for our studies because they are easy to play (even with very little experience) and are also easily modulated, in terms of game difficulty. However, there are a number of issues that may have stemmed from both the choice of game, and the determination of game difficulty. The first issue relating to the gameplay condition is that we could not verify the levels of difficulty that had been chosen. Although we attempted to create a variety of distinctly different difficulty levels, we have no way to ensure that each participant felt as though the level label matched the actual demand of the game. Game demand as a concept relates not only to the actual difficulty level of the game, but also to the skill level of the player. One way in which we could have potentially negated the effects of ubiquitous game demand would be via difficulty calibration. Had we calibrated the difficulty of the game for each participant (i.e. providing participants with a variety of game difficulty options and allowing them to rate their difficulty themselves), we may have been able to define more specific and participant-dependant game difficulty levels, which may have provided clearer results in terms of how game difficulty affected a participants response to pain. Another factor that may have affected participant's perception of game difficulty is that we did not account for previous gaming experience. As such, there is the potential that we had a wide variation of previous gaming experience within the participant pool, which could have also added unexpected variability to our results. Similarly, we did not address whether participants enjoyed racing games. This is important because if a participant did not enjoy such games, it would have been difficult for them to experience any of the benefits of immersion, which could have also skewed our results. All of these issues may have been present throughout the duration of our studies and could have also affected the levels of stimulation, motivation and attention that the game induced in participants as these factors could have influenced participant's perception and awareness of pain.

In relation to Study 3, there were a number of issues inherent with the creation of both the neuroadaptive and random-adaptive game, which may have prevented participants from experiencing the benefit of a neuroadaptive game and may have also obfuscated the results. Our hypothesis was that a neuroadaptive game would have a statistically significant impact on pain tolerance in comparison to a non-neuroadaptive game. Unfortunately, we were unable to confirm this hypothesis as statistical significance was not found for both the neuroadaptive and non-neuroadaptive game in Study 3. In terms of the neuroadaptive game, many of the same issues that were present in the standard Space Ribbon and Ribbon Rush games may still have been a contributing factor. Although the neuroadaptive game should have prevented the game from providing an unsuitable level of difficulty, the difficulty of the game was classified in real time using training data collected during Study 2. As previously mentioned, the game difficulty levels provided in Study 2 were not verified against participant's actual opinion on game difficulty. This could have meant that classification of game difficulty was not as efficient as it could have been. Equally, it is possible that racing games were not enjoyable for every participant. Issues that were more inherent with the neuroadaptive game itself relate to the time windows that were used for game adaptation. The neuroadaptive game had the ability to adapt game difficulty every 30 seconds. This window was selected to ensure that game difficulty did not modulate so rapidly that determining an optimal level of difficulty became a difficult task. However, on reflection, it is possible that this time period may have been too long, and therefore participants may have been playing the game at a level that was unsuitable for them for too long. it is postulated that reducing this time period may have led to better results in terms of the effectiveness of the neuroadaptive game.

Another issue that may have affected our results is the nature of the control condition. In order to determine whether a neuroadaptive game can provide a more immersive experience (measured via changes in pain tolerance) then there must be a comparison condition. In Study 3, a random adaptive game was created as our control condition. This game adapted at the same frequency as the neuroadaptive game, but would adjust the difficulty randomly, rather than based on the results of current game difficulty classification. Although this game would provide an efficient control condition, post-hoc analysis of game difficulty variation revealed that the random game and the neuroadaptive game were performing in very similar ways. As such, the two conditions may have been too similar to provide a clear indication of whether a neuroadaptive game would provide a better distraction from pain. On reflection, there are other options for the creation of a control condition that may have provided clearer results. One such condition we have considered is a yoked control, wherein the control condition would have been controlled by the neuroadaptive changes that were tailored to another participant. This game would have still been adaptive; however, it would not have been specifically tailored, which would indicate whether there is a statistically significant effect of pain tolerance between a neuroadaptive game and a non-neuroadaptive game. This technique has previously been used in neuroadaptive studies (Prinzel, Freeman, Scerbo, Mikulka, & Pope, 2003).

As such, many of the problems inherent in the research related to the game condition. However, there were also some issues with the fNIRS methodology. It was imperative for the creation of a neuroadaptive game that the pipeline for processing and classification of fNIRS data could perform in real time, which presented restrictions in terms of the types of filtration techniques that could be used. One such filter that may have provided clearer results is the Sliding Window Motion Artefact Rejection (SMAR) algorithm (Ayaz, Izzetoglu, Shewokis, & Onaral, 2010). This algorithm is designed to reduce the noise that can result from fNIRS optodes moving during data collection. It is possible that during data collection optodes can move slightly and lose connection with the scalp, which can cause artefacts in the data that may be misrepresented as an increase or decrease in oxygenated haemoglobin. The SMAR filter is appropriate for real-time use as it is reported to be a simple algorithm with appropriate run time and computational cost requirements. It is posited that the use of the SMAR filter may have improved the clarity of our data, which could, in turn, improve classification. Furthermore, the clarity of our data could have been further improved had we implemented spike detection algorithms, which are designed to locate and reduce spikes in the data that occur due to noise but may be misrepresentative of neuronal activity. Spike detection is frequently carried out post-hoc over a complete dataset, allowing for the determination of abrupt fluctuations within the data. However, many of these algorithms are not fit for purpose when considering a real-time application, as we simply did not have enough data at each timeframe to apply such filters. Applying spike detection algorithms over shorter windows could lead to the reduction in the representation of genuine neural activity (Cui, Bray, & Reiss, 2010b). One potential technique that could have been used in order to detect and remove spikes that are unrelated to true neural activity would be short-separation fNIRS channels, which measure extracerebral activity, including Meyer waves and blood pressure waves. The data collected from these channels can then be used to reduce the presence of noise within deep channels (actual data collection channels), which contain data relating to both cerebral and extracerebral activity (Goodwin, Gaudet, & Berger, 2014; Yücel et al., 2015). Data collected from the short separation channels can then be removed from the true fNIRS signal to ensure that the data is as clean as possible. It may have also been useful to use a device, such as the Artinis Polhemus, to precisely digitise the positions of the optodes on the participant's heads. Although the optode layout used was consistent for each study, across all participants, it is possible that mitigating factors, such as head shape and/or size, could have prevented the optodes from being placed in the correct position for each participant. Using a tool like a digitiser allows for certainty in optode placement, which in turn could improve data classification.

It has been noted that the approaches used for fNIRS data filtration can cause statistically significant differences between the same dataset. For example, a dataset that has been filtered using manufacturer-provided tools may be significantly different to the same dataset that has been treated using alternative filters recommended by the literature. It has been suggested that there are issues within the use of fNIRS technology that stem from the use (and lack of understanding) of manufacturer-built filtering algorithms. It has been suggested that fNIRS data should only be filtered using these techniques if the researcher is confident in the process that is undertaken to provide a filtered signal. Recommendations for overcoming these issues are partially outside of the scope of this project and require the determination of a set of standards that should be used for fNIRS data processing. However, there are suggestions, such as the use of regression techniques, which may provide a more accurately filtered signal (Pfeifer et al., 2018). Although we did not use any manufacturer provided software for the treatment of fNIRS data, we did use algorithms that had been designed specifically to reduce psychophysiological noise. One such method was the use of the CBSI algorithm, which was designed to reduce positively correlated data. However, this algorithm makes the assumption that all positively correlated data is related to a noise artefact and that HbO and HHb data have a constant ratio. If these assumptions are not correct, then the approach would not be valid (Brigadoi et al., 2014). It is also worth noting that, when the CBSI algorithm has been applied, transformed HHb data is derived mainly from HbO data rather than from the actual collected HHb data. This leads to the question of whether the CBSI algorithm is repairing issues within the correlation of these two data streams, or whether it is merely providing a reading of HHb, which is converse to that of HbO and so may not be an accurate assessment of the data. Further research is required to determine whether the CBSI algorithm is sufficient for use in fNIRS studies, or whether other avenues should be explored. One alternative option that could have been explored would be to design the actual data-processing loop in a way that simply ignored large spikes in the data, rather than attempting to filter them from the signal, and prevent classification from occurring over data that had a large deviation from the previously collected dataset.

Data classification was, as previously mentioned, carried out using SI classification. We believe that this is the most efficient and effective way to build a real-time neuroadaptive system. However, it could be hypothesised that using SD classification may have provided more accurate game difficulty modulation for each individual participant. It is possible that the most effective way to create a neuroadaptive system would be using SD classification – however, we believe that the benefits of SI classification outweigh the drawbacks. We believe that other avenues (the use of an alternation control condition, changes to the fNIRS system) should be prioritised over differing classification techniques. Only after alternate adaptation to the study have been carried out and reviewed would we consider moving to SD classification – if other previously mentioned recommendations were still ineffective in exploring our hypothesis.

Overall, there were issues inherent within our study relating to both the creation of the neuroadaptive game (and its control counterpart) and the actual methods that were used for fNIRS filtration. If these studies were to be replicated or expanded upon, we would consider changes to both the data processing pipeline and to the creation of the gaming conditions. Despite the limitations discussed, we believe the results that were gathered were promising and the limitations have provided interesting potential for future work.

7.4 Future work

There are many avenues that could be explored to further this research, such as creating a system that is more sensitive to determining the level of game demand. As discussed in the limitations of this chapter, altering the time windows for game difficulty modulation may improve the neuroadaptive game. However, using shorter time windows is likely to reduce classification accuracy, as there is simply a smaller amount of data available for classification. One way in which both the classification accuracy and the effectiveness of the neuroadaptive game could be improved is to reduce the time window of adaptation, but use a combination of fNIRS and Electroencephalogram (EEG) to attempt to improve or maintain classification accuracy, even in a shorter time window (Buccino, Keles, & Omurtag, 2016; Hirshfield et al., 2009; Leamy, Collins, & Ward, 2011). Another way classification accuracy of the data (and therefore the effectiveness of the neuroadaptive game) could be improved would be to employ subject-dependent classification techniques. In such a case, participants would experience a variety of game difficulty levels and then rate their engagement with each level. The data collected via these trials would then be labelled based on the participant's perception of game difficulty and/or engagement. Moving forward, the participant could then play a neuroadaptive game that is designed specifically to adapt to their personal perceptions of game difficulty rather than the current subject-independent technique that was used. It is postulated that this may mitigate issues relating to personal gaming experience and enjoyment.

For this research project, driving games were chosen because they are easy to play, even as a novice, there are a limited amount of controls to master and the overall goal of the game is explicit and unchanging. However, although racing games can be immersive, the nature of the games themselves may have highlighted some issues in our studies. One such way in which a racing game may not have been the most efficient choice is that racing games (or any game with a competitive element) are particularly stimulating. Therefore, although the game itself may be distracting, these games may raise the heart rate (HR) and blood pressure (BP) of a participant. An increase in HR and BP can lead to a heighted experience of pain, which is the opposite effect that is intended. Games with an element of competition can also elicit negative emotions (e.g. anger, sadness etc.), which are also known to increase the experience of pain, in some cases. Using a different style of game, possibly a game that has no competitive element, may have provided a similar level of immersion to the participant but without potentially inducing negative emotions, which may affect the participant's pain tolerance or experience.

Another interesting avenue to explore would be the use of thermal pain. Using thermal pain, as opposed to the CPT, provides the overall benefit that pain can be induced at any point, without requiring action from the participant. As such, requiring the participant to place their foot into the CPT may have disrupted potential experiences of immersion and caused participants to attend to the pain. Additionally, if pain is induced by thermal means then it could be controlled by the experimenter. Another benefit of using thermal pain is that the induction of pain could be based on a participant's perception of immersion. For example, during Study 3, all participants placed their foot into the CPT at the same stage of the game, regardless of the current level of game difficulty. As such, this could have meant that some participants were actually not immersed at this point. However, if an algorithm were created that controlled the thermal pain stimuli (and therefore only allowed the participant to experience pain when they were immersed in the distraction condition), this may provide a greater benefit to the neuroadaptive game.

The overall aim of this project was to attempt to reduce pain via neuroadaptive gaming. Although we were unable to find a correlation between the neuroadaptive game and a reduction in the pain experience, there are still many avenues to explore. One such element that we are keen to investigate further is the use of a neuroadaptive game in a clinical setting. However, there are a variety of factors that must be considered before this could be a feasible possibility. Infection control is one element that would be particularly pressing. As of 2018, it was estimated that there were around 300,000 healthcare associated infections annually[10].Bringing neuroadaptive gaming to a clinical setting would require the introduction of technology that is not typically found in a hospital setting (such as fNIRS equipment). The most pressing issue when considering infection control is the sensors and equipment that the patient would come into direct contact with. It would be imperative in this case that the required

[10] https://commonslibrary.parliament.uk/research-briefings/cdp-2018-0116/

equipment was reduced as much as possible. For this reason, it would be pertinent to determine (prior to the use of such an application in a clinical setting) the most relevant fNIRS data channels for classification. Determining these channels would then reduce the amount of optodes that would be required, which may mitigate some infection potential. We would also have to ensure, for both infection control purposes and ease of use, that the console on which the neuroadaptive game was to be played is portable and no larger than required (for ease of sterilisation). For this reason, it would be interesting to consider playing the neuroadaptive game on a handheld console, but considerations must then also be made as to how console and/or screen size may affect immersion (Hou, Nam, Peng, & Lee, 2012; Thompson, Nordin, & Cairns, 2012).

Although this project has been focused on the reduction of pain via neuroadaptive gaming, the pipeline for neuroadaptation could be used in a variety of settings. The backbone of the neuroadaptive loop that has been created is based on task engagement. An assumption, based on previous literature, was that a more difficult game would provide a more engaging experience. This means that any task that requires a level of engagement could potentially benefit from the employment of a neuroadaptive loop. This loop could then be used in tasks (real-world or otherwise) where there is an element of shared brain-machine control. For example, there are a variety of autopilot features that are used in both driving and flying situations, but the use of these autopilot systems can sometimes lead to operator disengagement (Gouraud, Delorme, & Berberian, 2017; Pope et al., 1995; Verdière et al., 2018). Disengagement can lead to accidents occurring in the same way that overload leaves less attention available for dealing with unexpected situations. Therefore, if a neuroadaptive system was created that could control when certain autopilot functions could be used (for example, decreasing task load by engaging autopilot when demand is high, and increasing task load by disengaging autopilot when task engagement is low) this could lead to safer and more efficient control over such vehicles. Overall, there are improvements that could be made to the current studies in order to thoroughly test the hypothesis, as well as many avenues that could be explored to further the scope of the research.

References

Abdelmoniem, S. A., & Mahmoud, S. A. (2015). Comparative Evaluation of Passive, Active, and Passive-Active Distraction techniques on Pain Perception during Local Anesthesia Administration in Children. *Journal of Advanced Research*, *7*(3), 551–556. https://doi.org/10.1016/j.jare.2015.10.001

Adams, N., Little, T. D., & Ryan, R. M. (2017). Self-determination theory. *Development of Self-Determination Through the Life-Course*, *55*(1), 47–54. https://doi.org/10.1007/978-94-024-1042-6_4

Akizuki, K., & Ohashi, Y. (2015). Measurement of functional task difficulty during motor learning: What level of difficulty corresponds to the optimal challenge point? *Human Movement Science*, *43*, 107–117. https://doi.org/10.1016/j.humov.2015.07.007

Anand, K. J. S., Willson, D. F., Berger, J., Harrison, R., Meert, K. L., Zimmerman, J., … Nicholson, C. (2010). Tolerance and Withdrawal From Prolonged Opioid Use in Critically Ill Children. *Pediatrics*, *125*(5), e1208–e1225. https://doi.org/10.1542/peds.2009-0489

Andreessen, L. M., Gerjets, P., Meurers, D., & Zander, T. O. (2020). Toward neuroadaptive support technologies for improving digital reading: a passive BCI-based assessment of mental workload imposed by text difficulty and presentation speed during reading. *User Modeling and User-Adapted Interaction*. https://doi.org/10.1007/s11257-020-09273-5

Arrington, C. M., Carr, T. H., Mayer, A. R., & Rao, S. M. (2000). Neural Mechanisms of Visual Attention: Object-Based Selection of a Region in Space. *Journal of Cognitive Neuroscience*, *12*(supplement 2), 106–117. https://doi.org/10.1162/089892900563975

Auria, L., & Moro, R. A. (2011). Support Vector Machines (SVM) as a Technique for Solvency Analysis. *SSRN Electronic Journal*, (August). https://doi.org/10.2139/ssrn.1424949

Ayaz, H., Izzetoglu, M., Shewokis, P. A., & Onaral, B. (2010). Sliding-window motion artifact rejection for functional Near-Infrared Spectroscopy. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*, 6567–6570. https://doi.org/10.1109/IEMBS.2010.5627113

Ayaz, H., Onaral, B., Izzetoglu, K., Shewokis, P. A., McKendrick, R., & Parasuraman, R. (2013). Continuous monitoring of brain dynamics with functional near infrared spectroscopy as a tool for neuroergonomic research: empirical examples and a technological development. *Frontiers in Human Neuroscience*, *7*(December), 871. https://doi.org/10.3389/fnhum.2013.00871

Baker, W. B., Parthasarathy, A. B., Busch, D. R., Mesquita, R. C., Greenberg, J. H., & Yodh, A. G. (2014). Modified Beer-Lambert law for blood flow. *Biomedical Optics Express*, *5*(11), 4053. https://doi.org/10.1364/BOE.5.004053

Bantick, S. J., Wise, R. G., Ploghaus, A., Clare, S., Smith, S. M., & Tracey, I. (2002). Imaging how attention modulates pain in humans using functional MRI. *Brain : A Journal of Neurology*, *125*(Pt 2), 310–319. https://doi.org/10.1093/brain/awf022

Bastos, A. M., & Schoffelen, J. M. (2016). A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Frontiers in Systems Neuroscience*, *9*(JAN2016), 1–23. https://doi.org/10.3389/fnsys.2015.00175

Bauernfeind, G., Wriessnegger, S. C., Daly, I., & Müller-Putz, G. R. (2014). Separating heart and brain: on the reduction of physiological noise from multichannel functional near-infrared spectroscopy (fNIRS) signals. *Journal of Neural Engineering*, *11*(5), 056010. https://doi.org/10.1088/1741-2560/11/5/056010

Berger, A., Horst, F., Müller, S., Steinberg, F., & Doppelmayr, M. (2019). Current state and future prospects of EEG and fNIRS in robot-assisted gait rehabilitation: A brief review. *Frontiers in Human Neuroscience*, *13*(June). https://doi.org/10.3389/fnhum.2019.00172

Bontchev, B. (2016). Adaptation in affective video games: A literature review. *Cybernetics and Information Technologies*, *16*(3), 3–34. https://doi.org/10.1515/cait-2016-0032

Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., … Cooper, R. J. (2014). Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data. *NeuroImage*, *85*, 181–191. https://doi.org/10.1016/j.neuroimage.2013.04.082

Broad, R. D., & Wheeler, K. (2006). An adult with childhood medical trauma treated with psychoanalytic psychotherapy and EMDR: A case study. *Perspectives in Psychiatric Care*, *42*(2), 95–105. https://doi.org/10.1111/j.1744-6163.2006.00058.x

Broadbent, D. E. (1956). Successive Responses to Simultaneous Stimuli. *Quarterly Journal of Experimental Psychology*, *8*(4), 145–152. https://doi.org/10.1080/17470215608416814

Buccino, A. P., Keles, H. O., & Omurtag, A. (2016). Hybrid EEG-fNIRS asynchronous brain-computer interface for multiple motor tasks. *PLoS ONE*, *11*(1), 1–16. https://doi.org/10.1371/journal.pone.0146610

Bullinger, M., Naber, D., Pickar, D., Cohen, R. M., Kalin, N. H., Pert, A., & Bunney, W. E. (1984). Endocrine effects of the cold pressor test: Relationships to subjective pain appraisal and coping. *Psychiatry Research*, *12*(3), 227–233. https://doi.org/10.1016/0165-1781(84)90028-3

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, Vol. 97, pp. 523–547. https://doi.org/10.1037/0033-295X.97.4.523

Cairns, P., Cox, A., & Imran Nordin, A. (2014). Immersion in Digital Games: Review of Gaming Experience Research. *Handbook of Digital Games*, 339–361. https://doi.org/10.1002/9781118796443.ch12

Chanel, G., Rebetez, C., Bétrancourt, M., & Pun, T. (2008). Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. *Proceedings of the 12th International Conference on Entertainment and Media in the Ubiquitous Era - MindTrek '08*, (August), 13. https://doi.org/10.1145/1457199.1457203

Chanel, G., Rebetez, C., Bétrancourt, M., & Pun, T. (2011). *Emotion Assessment From Physiological Signals for Adaptation of Game Difficulty*. *41*(6), 1052–1063.

City, R. (2017). *(19) United States (12)* (Vol. 1).

Cooper, T. E., Fisher, E., Gray, A. L., Krane, E., Sethna, N., van Tilburg, M. A. L., … Wiffen, P. J. (2017). Opioids for chronic non-cancer pain in children and adolescents. *Cochrane Database of Systematic Reviews*, *2017*(2). https://doi.org/10.1002/14651858.CD012538

Cope, M. (1991). The development of a near infrared spectroscopy system and its application for non invasive monitory of cerebral blood and tissue oxygenation in the newborn infants. *Diss. University of London*, 342. Retrieved from http://discovery.ucl.ac.uk/1317956/

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Erratum to "Voluntary orienting is dissociated from target detection in human posterior parietal cortex." *Nature Neuroscience*, *3*(5), 521–521. https://doi.org/10.1038/74905

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Coyle, S. M., Ward, T. E., Markham, C. M., Ward, E., & Markham, C. M. (2007). Brain-computer interface using a simplified functional near-infrared spectroscopy system. *Journal of Neural Engineering*, *4*(3), 219–226. https://doi.org/10.1088/1741-2560/4/3/007

Crombez, G., Van Ryckeghem, D. M. L., Eccleston, C., & Van Damme, S. (2013). Attentional bias to pain-related information: A meta-analysis. *Pain*, *154*(4), 497–510. https://doi.org/10.1016/j.pain.2012.11.013

Csikszentmihalyi, M. (1975). Beyond Boredom and Anxiety: Experiencing Flow in Work and Play. *The Jossey-Bass Behavioral Science Series*, 231. https://doi.org/10.2307/2065805

Cui, X., Bray, S., & Reiss, A. L. (2010a). Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *NeuroImage*, *49*(4), 3039–3046. https://doi.org/10.1016/j.neuroimage.2009.11.050

Cui, X., Bray, S., & Reiss, A. L. (2010b). Speeded near infrared spectroscopy (NIRS) response detection. *PLoS ONE*, *5*(11). https://doi.org/10.1371/journal.pone.0015474

Cui, X., Bryant, D. M., & Reiss, A. L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *NeuroImage*, *59*(3), 2430–2437. https://doi.org/10.1016/j.neuroimage.2011.09.003

Czub, M., & Piskorz, J. (2014). How body movement influences virtual reality analgesia? *Proceedings - 2014 International Conference on Interactive Technologies and Games, ITAG 2014*, 13–19. https://doi.org/10.1109/iTAG.2014.8

Dehais, F., Causse, M., Vachon, F., Régis, N., Menant, E., & Tremblay, S. (2014). Failure to detect critical auditory alerts in the cockpit: Evidence for inattentional deafness. In *Human Factors* (Vol. 56). https://doi.org/10.1177/0018720813510735

Diseth, T. H. (2006). Dissociation following traumatic medical treatment procedures in childhood: A longitudinal follow-up. *Development and Psychopathology*, *18*(1), 233–251. https://doi.org/10.1017/S0954579406060135

Durantin, G., Dehais, F., Gonthier, N., Terzibas, C., & Callan, D. E. (2017). Neural signature of inattentional deafness. *Human Brain Mapping*, *38*(11), 5440–5455. https://doi.org/10.1002/hbm.23735

Ebrahimpour, F., Sadeghi, N., Najafi, M., Iraj, B., & Shahrokhi, A. (2015). Effect of Playing Interactive Computer Game on Distress of Insulin Injection Among Type 1 Diabetic Children. *Iranian Journal of Pediatrics*, *25*(3), e427. https://doi.org/10.5812/ijp.25(3)2015.427

Eccleston, C., & Crombez, G. (1999). Pain demands attention: a cognitive-affective model of the interruptive function of pain. *Psychol Bull*, *125*(3), 356–366. https://doi.org/10.1037/0033-2909.125.3.356

Ewing, K. C., Fairclough, S. H., & Gilleade, K. (2016). Evaluation of an Adaptive Game that Uses EEG Measures Validated during the Design Process as Inputs to a Biocybernetic Loop. *Frontiers in Human Neuroscience*, *10*(May), 1–13. https://doi.org/10.3389/fnhum.2016.00223

Fairclough, S., & Gilleade, K. (2012a). Construction of the Biocybernetic Loop: A Case Study. *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, *Santa Moni*, 571–578. https://doi.org/10.1145/2388676.2388797

Fairclough, S., & Gilleade, K. (2012b). Construction of the biocybernetic loop. *Proceedings of the 14th ACM International Conference on Multimodal Interaction - ICMI '12*, 571. https://doi.org/10.1145/2388676.2388797

Fairclough, S. H. (2009a). Fundamentals of physiological computing. *Interacting with Computers*, *21*(1–2), 133–145. https://doi.org/10.1016/j.intcom.2008.10.011

Fairclough, S. H. (2009b). Fundamentals of physiological computing. *Interacting with Computers*, *21*(1–2), 133–145. https://doi.org/10.1016/j.intcom.2008.10.011

Fairclough, S. H. (2010). *Physiological Computing: Interfacing with the Human Nervous System*. 1–20. https://doi.org/10.1007/978-90-481-3258-4_1

Fairclough, S. H., Gilleade, K., Ewing, K. C., & Roberts, J. (2013a). Capturing user engagement via psychophysiology: measures and mechanisms for biocybernetic adaptation. *International Journal of Autonomous and Adaptive Communications Systems*, *6*(1), 63. https://doi.org/10.1504/IJAACS.2013.050694

Fairclough, S. H., Gilleade, K., Ewing, K. C., & Roberts, J. (2013b). Capturing user engagement via psychophysiology: Measures and mechanisms for biocybernetic adaptation. *International Journal of Autonomous and Adaptive Communications Systems*, *6*(1), 63–79. https://doi.org/10.1504/IJAACS.2013.050694

Fairclough, S. H., Stamp, K., Dobbins, C., & Poole, H. M. (2020). Computer games as distraction from PAIN: Effects of hardware and difficulty on pain tolerance and subjective IMMERSION. *International Journal of Human Computer Studies*, *139*(May 2019). https://doi.org/10.1016/j.ijhcs.2020.102427

Fernández, H. D., Koji, M., & Kondo, K. (2017). Adaptable game experience based on player's performance and EEG. *Proceedings - 2017 NICOGRAPH International, NICOInt 2017*, (June), 1–8. https://doi.org/10.1109/NICOInt.2017.11

Ferrari, M., & Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage*, *63*(2), 921–935. https://doi.org/10.1016/j.neuroimage.2012.03.049

Fishburn, F. a, Norr, M. E., Medvedev, A. V, & Vaidya, C. J. (2014). Sensitivity of fNIRS to cognitive state and load. *Frontiers in Human Neuroscience*, *8*(February), 76. https://doi.org/10.3389/fnhum.2014.00076

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National*

Academy of Sciences of the United States of America, *103*(26), 10046–10051. https://doi.org/10.1073/pnas.0604187103

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(27), 9673–9678. https://doi.org/10.1073/pnas.0504136102

Freeman, F. G., Mikulka, P. J., Prinzel, L. J., & Scerbo, M. W. (1999). Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology*, *50*(1), 61–76. https://doi.org/10.1016/S0301-0511(99)00002-2

Girouard, A., Solovey, E. T., Hirshfield, L. M., Chauncey, K., Sassaroli, A., Fantini, S., & Jacob, R. J. K. (2009). Distinguishing Difficulty Levels with Non-invasive Brain. *Proceedings of INTERACT 2009*, (Figure 1), 440–452.

Goldman, A., & Lloyd-Thomas, A. R. (2004). Pain management in children. *Current Anaesthesia and Critical Care*, *15*(4–5), 272–283. https://doi.org/10.1016/j.cacc.2004.08.014

Goodwin, J. R., Gaudet, C. R., & Berger, A. J. (2014). Short-channel functional near-infrared spectroscopy regressions improve when source-detector separation is reduced. *Neurophotonics*, *1*(1), 015002. https://doi.org/10.1117/1.nph.1.1.015002

Gordon, N. S., Merchant, J., Zanbaka, C., Hodges, L. F., & Goolkasian, P. (2011). Interactive gaming reduces experimental pain with or without a head mounted display. *Computers in Human Behavior*, *27*(6), 2123–2128. https://doi.org/10.1016/j.chb.2011.06.006

Gould, D., Goldstone, L., Kelly, D., & Gammon, J. (2004). Examining the validity of pressure ulcer risk assessment scales: A replication study. *International Journal of Nursing Studies*, *41*(3), 331–339. https://doi.org/10.1016/j.ijnurstu.2003.10.005

Gouraud, J., Delorme, A., & Berberian, B. (2017). Autopilot, mind wandering, and the out of the loop performance problem. *Frontiers in Neuroscience*, *11*(OCT), 1–13. https://doi.org/10.3389/fnins.2017.00541

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, *11*(5/6), 561–566. https://doi.org/10.5194/npg-11-561-2004

Guadagnoll, M. A., & Lee, T. D. (2004). Challenge Point: A Framework for Conceptualizing the Effects of Various Practice Conditions in Motor Learning. *Journal of Motor Behavior*, *36*(2), 212–224. https://doi.org/10.3200/JMBR.36.2.212-224

Gupta, D., Agarwal, A., Dhiraaj, S., Tandon, M., Kumar, M., Singh, R. S., … Singh, U. (2006). An evaluation of efficacy of balloon inflation on venous cannulation pain in children: A prospective, randomized, controlled study. *Anesthesia and Analgesia*, *102*(5), 1372–1375. https://doi.org/10.1213/01.ane.0000205741.82299.d6

Gwinup, G., Haw, T., & Elias, A. (1983). Cardiovascular changes in video-game players. *Postgraduate Medicine*, *74*(6), 245–248. https://doi.org/10.1080/00325481.1983.11698546

Harrivel, A. R., Weissman, D. H., Noll, D. C., & Peltier, S. J. (2013). Monitoring attentional state with fNIRS. *Frontiers in Human Neuroscience*, *7*(December), 861.

https://doi.org/10.3389/fnhum.2013.00861
Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, *52*, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9
Heger, D., Herff, C., Putze, F., Mutter, R., & Schultz, T. (2014). Continuous Affective States Recognition using functional Near Infrared Spectroscopy. *Brain Computer Interfaces*. https://doi.org/DOI 10.1080/2326263X.2014.912884
Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F., & Schultz, T. (2014). Mental workload during n-back task—quantified in the prefrontal cortex using fNIRS. *Frontiers in Human Neuroscience*, *7*(January), 1–9. https://doi.org/10.3389/fnhum.2013.00935
Hirshfield, L. M., Chauncey, K., Gulotta, R., Girouard, A., Solovey, E. T., Jacob, R. J. K., … Fantini, S. (2009). Combining Electroencephalograph and Near Infrared Spectroscopy to Explore Users' Instantaneous and Continuous Mental Workload States. *HCI International 2009 13th International Conference on Human-Computer Interaction*, 239–247. https://doi.org/10.1007/978-3-642-02812-0_28
Hoffman, H. G. (2004). Virtual-reality therapy. *Scientific American*, *291*(2), 58–65. https://doi.org/10.1038/scientificamerican0804-58
Hoffman, H. G., Doctor, J. N., Patterson, D. R., Carrougher, G. J., & Furness, T. A. (2000). Virtual reality as an adjunctive pain control during burn wound care in adolescent patients. *Pain*, *85*(1–2), 305–309. https://doi.org/10.1016/S0304-3959(99)00275-4
Hou, J., Nam, Y., Peng, W., & Lee, K. M. (2012). Effects of screen size, viewing angle, and players' immersion tendencies on game experience. *Computers in Human Behavior*, *28*(2), 617–623. https://doi.org/10.1016/j.chb.2011.11.007
Hunicke, R. (2005). The case for dynamic difficulty adjustment in games. *ACM International Conference Proceeding Series*, *265*(January 2005), 429–433. https://doi.org/10.1145/1178477.1178573
Hussein, H. A. (2015). Effect of active and passive distraction on decreasing pain associated with painful medical procedures among school aged children. *World J Nurs Sci*, *1*(2), 13–23. https://doi.org/10.5829/idosi.wjns.2015.1.2.93202
Inan, G., & Inal, S. (2019). The impact of 3 different distraction techniques on the pain and anxiety levels of children during venipuncture. *Clinical Journal of Pain*, *35*(2), 140–147. https://doi.org/10.1097/AJP.0000000000000666
Jameson, E., Trevena, J., & Swain, N. (2011a). Electronic gaming as pain distraction. *Pain Research and Management*, *16*(1), 27–32. https://doi.org/10.1155/2011/856014
Jameson, E., Trevena, J., & Swain, N. (2011b). Electronic gaming as pain distraction. *Pain Research and Management*, *16*(1), 27–32.
Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human Computer Studies*, *66*(9), 641–661. https://doi.org/10.1016/j.ijhcs.2008.04.004
Jennings-Teats, M., Smith, G., & Wardrip-Fruin, N. (2010). Polymorph: Dynamic Difficulty Adjustment

through level generation. *Workshop on Procedural Content Generation in Games, PC Games 2010, Co-Located with the 5th International Conference on the Foundations of Digital Games*, (September), 2–6. https://doi.org/10.1145/1814256.1814267

Johnson, M. H. (2005). How does distraction work in the management of pain? *Current Pain and Headache Reports*, *9*(2), 90–95. https://doi.org/10.1007/s11916-005-0044-1

Julien, C. (2006). The enigma of Mayer waves: Facts and models. *Cardiovascular Research*, *70*(1), 12–21. https://doi.org/10.1016/j.cardiores.2005.11.008

Kaheni, S., Rezai, M. S., Bagheri-nesami, M., & Hossein, A. (n.d.). *The Effect of Distraction Technique on the Pain of Dressing Change among 3-6 Year-old Children*. *4*(28), 1603–1610.

Kamran, M. A., Mannann, M. M. N., & Jeong, M. Y. (2018). Differential path-length factor's effect on the characterization of brain's hemodynamic response function: A functional near-infrared study. *Frontiers in Neuroinformatics*, *12*(June), 1–15. https://doi.org/10.3389/fninf.2018.00037

Keogh, E., Moore, D. J., Duggan, G. B., Payne, S. J., & Eccleston, C. (2013). The disruptive effects of pain on complex cognitive performance and executive control. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0083272

Kincade, J. M., Abrams, R. A., Astafiev, S. V., Shulman, G. L., & Corbetta, M. (2005). An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *Journal of Neuroscience*, *25*(18), 4593–4604. https://doi.org/10.1523/JNEUROSCI.0236-05.2005

Koenraadt, K. L. M., Roelofsen, E. G. J., Duysens, J., & Keijsers, N. L. W. (2014). Cortical control of normal gait and precision stepping: An fNIRS study. *NeuroImage*, *85*, 415–422. https://doi.org/10.1016/j.neuroimage.2013.04.070

Kononenko, I., Šimec, E., & Robnik-Šikonja, M. (1997). Overcoming the myopia of inductive learning algorithms with RELIEFF. *Applied Intelligence*, *7*(1), 39–55. https://doi.org/10.1023/A:1008280620621

Krol, L. R., Freytag, S. C., & Zander, T. O. (2017). Meyendtris: A hands-free, multimodal tetris clone using eye tracking and passive BCI for intuitive neuroadaptive gaming. *ICMI 2017 - Proceedings of the 19th ACM International Conference on Multimodal Interaction*, *2017-Janua*, 433–437. https://doi.org/10.1145/3136755.3136805

Laird, A. R., Eickhoff, S. B., Rottschy, C., Bzdok, D., Ray, K. L., & Fox, P. T. (2013). Networks of task co-activations. *NeuroImage*, *80*, 505–514. https://doi.org/10.1016/j.neuroimage.2013.04.073

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N., Beck, D. M., & Konstantinou, N. (2014). Blinded by the load: Attention, awareness and the role of perceptual load. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1641). https://doi.org/10.1098/rstb.2013.0205

Law, Dahlquist, Still, Weiss, Herbert, Wohlheiter, & S, H. (2011). Videogame distraction using virtual reality technology for children exeperiencing cold pressor pain: the role of cognitive processing. *Journal of Pediatric Psychology*, *36*(1), 84–94.

Leamy, D. J., Collins, R., & Ward, T. E. (2011). Combining fNIRS and EEG to improve motor cortex activity classification during an imagined movement-based task. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6780 LNAI*, 177–185. https://doi.org/10.1007/978-3-642-21852-1_22

Legrain, V., Damme, S. Van, Eccleston, C., Davis, K. D., Seminowicz, D. A., & Crombez, G. (2009). A neurocognitive model of attention to pain: Behavioral and neuroimaging evidence. *Pain*, *144*(3), 230–232. https://doi.org/10.1016/j.pain.2009.03.020

Liu, C., Agrawal, P., Sarkar, N., & Chen, S. (2009). Dynamic Difficulty Adjustment in Computer Games Through Real-Time Anxiety-Based Affective Feedback. *International Journal of Human-Computer Interaction*, *25*(6), 506–529. https://doi.org/10.1080/10447310902963944

Liu, Y., Piazza, E. A., Simony, E., Shewokis, P. A., Onaral, B., Hasson, U., & Ayaz, H. (2017). Measuring speaker-listener neural coupling with functional near infrared spectroscopy. *Scientific Reports*, *7*(February), 1–13. https://doi.org/10.1038/srep43293

Lovallo, W. (1975). The Cold Pressor Test and Autonomic Function: A Review and Integration. *Psychophysiology*, *12*(3), 268–282. https://doi.org/10.1111/j.1469-8986.1975.tb01289.x

MacLaren, J. E., & Cohen, L. L. (2005). A comparison of distraction strategies for venipuncture distress in children. *Journal of Pediatric Psychology*, *30*(5), 387–396. https://doi.org/10.1093/jpepsy/jsi062

Matson, K. L., Johnson, P. N., Tran, V., Horton, E. R., & Sterner-Allison, J. (2019). Opioid use in children. *Journal of Pediatric Pharmacology and Therapeutics*, *24*(1), 72–75. https://doi.org/10.5863/1551-6776-24.1.72

Matthews, G., Joyner, L., Gilliand, K., Campbell, S., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire- Towards a state "big three"? *Personality Psychology*, *7*, 335–350.

Mayer, A. R., Toulouse, T., Klimaj, S., Ling, J. M., Pena, A., & Bellgowan, P. S. F. (2014). Investigating the Properties of the Hemodynamic Response Function after Mild Traumatic Brain Injury. *Journal of Neurotrauma*, *31*(2), 189–197. https://doi.org/10.1089/neu.2013.3069

Meng, L., Pei, G., Zheng, J., & Ma, Q. (2016). Close games versus blowouts: Optimal challenge reinforces one's intrinsic motivation to win. *International Journal of Psychophysiology*, *110*, 102–108. https://doi.org/10.1016/j.ijpsycho.2016.11.001

Metz, A. J., Wolf, M., Achermann, P., & Scholkmann, F. (2015). A new approach for automatic removal of movement artifacts in near-infrared spectroscopy time series by means of acceleration data. *Algorithms*, *8*(4), 1052–1075. https://doi.org/10.3390/a8041052

Miller, K., Rodger, S., Bucolo, S., Greer, R., & Kimble, R. M. (2010). Multi-modal distraction. Using technology to combat pain in young children with burn injuries. *Burns*, *36*(5), 647–658. https://doi.org/10.1016/j.burns.2009.06.199

Mitchell, L. A., MacDonald, R. A. R., & Brodie, E. E. (2004). Temperature and the cold pressor test. *The Journal of Pain*, *5*(4), 233–237. https://doi.org/10.1016/j.jpain.2004.03.004

Muñoz, J. E., Gouveia, E. R., Cameirão, M. S., & Badia, S. B. I. (2017). The biocybernetic loop engine: An integrated tool for creating physiologically adaptive videogames. *PhyCS 2017 - Proceedings of the 4th International Conference on Physiological Computing Systems*, (PhyCS), 45–54. https://doi.org/10.5220/0006429800450054

Nacke, L. E., & Lindley, C. A. (2008). Flow and Immersion in First-person Shooters: Measuring the Player's Gameplay Experience. *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, 81–88. https://doi.org/10.1145/1496984.1496998

Naito, M., Michioka, Y., Ozawa, K., Ito, Y. I., Kiguchi, M., & Kanazawa, T. (2007). A communication means for totally locked-in ALS patients based on changes in cerebral blood volume measured with near-infrared light. *IEICE Transactions on Information and Systems*, *E90-D*(7), 1028–1037. https://doi.org/10.1093/ietisy/e90-d.7.1028

Naseer, N., & Hong, K.-S. (2015). fNIRS-based brain-computer interfaces: a review. *Frontiers in Human Neuroscience*, *9*(January), 1–15. https://doi.org/10.3389/fnhum.2015.00003

Nilsson, S., Enskär, K., Hallqvist, C., & Kokinsky, E. (2013). Active and Passive Distraction in Children Undergoing Wound Dressings. *Journal of Pediatric Nursing*, *28*(2), 158–166. https://doi.org/10.1016/j.pedn.2012.06.003

Norman, D. A., & Shallice, T. (1980). *Attention to action: Willed and automatic control of behavior. Report No 8006.* (January 1986), 43. https://doi.org/10.1007/978-1-4757-0629-1_1

Nouwen, A., Cloutier, C., Kappas, A., Warbrick, T., & Sheffield, D. (2006). Effects of Focusing and Distraction on Cold Pressor–Induced Pain in Chronic Back Pain Patients and Control Subjects. *The Journal of Pain*, *7*(1), 62–71. https://doi.org/10.1016/j.jpain.2005.08.004

Parker, M., Delahunty, B., Heberlein, N., Devenish, N., Wood, F. M., Jackson, T., … Edgar, D. W. (2016). Interactive gaming consoles reduced pain during acute minor burn rehabilitation: A randomized, pilot trial. *Burns*, *42*(1), 91–96. https://doi.org/10.1016/j.burns.2015.06.022

Pfeifer, M. D., Scholkmann, F., & Labruyère, R. (2018). Signal Processing in Functional Near-Infrared Spectroscopy (fNIRS): Methodological Differences Lead to Different Statistical Results. *Frontiers in Human Neuroscience*, *11*(January), 1–12. https://doi.org/10.3389/fnhum.2017.00641

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, *40*(1–2), 187–195. https://doi.org/10.1016/0301-0511(95)05116-3

Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2000). A Closed-Loop System for Examining Psychophysiological Measures for Adaptive Task Allocation. *The International Journal of Aviation Psychology*, *10*(4), 393–410. https://doi.org/10.1207/S15327108IJAP1004_6

Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2003). Effects of a Psychophysiological System for Adaptive Automation on Performance, Workload, and the Event-Related Potential P300 Component. *Human Factors*, *45*(4), 601–613. https://doi.org/10.1518/hfes.45.4.601.27092

Przybylski, A. K., Deci, E. L., Rigby, C. S., & Ryan, R. M. (2014). Competence-Impeding Electronic Games and Players' Aggressive Feelings, Thoughts, and Behaviors. *Journal of Personality and Social Psychology*, *106*(3), 441–457. https://doi.org/10.1037/a0034820

Racz, F. S., Mukli, P., Nagy, Z., & Eke, A. (2017). Increased prefrontal cortex connectivity during cognitive challenge assessed by fNIRS imaging. *Biomedical Optics Express*, *8*(8), 3842.

https://doi.org/10.1364/boe.8.003842

Rémy, F., Frankenstein, U. N., Mincic, A., Tomanek, B., & Stroman, P. W. (2003). Pain modulates cerebral activity during cognitive performance. *NeuroImage*, *19*(3), 655–664. https://doi.org/10.1016/S1053-8119(03)00146-0

Rheinberg, F. (2008). Intrinsic Motivation and Flow. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation and Action* (pp. 323–348). https://doi.org/10.1017/CBO9780511499821.014

Rutter, C. E., Dahlquist, L. M., & Weiss, K. E. (2009). Sustained Efficacy of Virtual Reality Distraction. *The Journal of Pain*, *10*(4), 391–397. https://doi.org/10.1016/j.jpain.2008.09.016

Scalf, P. E., Torralbo, A., Tapia, E., & Beck, D. M. (2013). Competition explains limited attention and perceptual resources: Implications for perceptual load and dilution theories. *Frontiers in Psychology*, *4*(MAY), 1–9. https://doi.org/10.3389/fpsyg.2013.00243

Scerbo, M. W., Freeman, F. G., & Mikulka, P. J. (2003). A brain-based system for adaptive automation. *Theoretical Issues in Ergonomics Science*, *4*(1–2), 200–219. https://doi.org/10.1080/1463922021000020891

Scholkmann, F., & Wolf, M. (2013). General equation for the differential pathlength factor of the frontal human head depending on wavelength and age. *Journal of Biomedical Optics*, *18*(10), 105004. https://doi.org/10.1117/1.jbo.18.10.105004

Selesnick, I. W., & Burrus, C. S. (1998). *H 1 = h 2 = a=d* . *46*(6), 1688–1694.

Sitaram, R., Zhang, H., Guan, C., Thulasidas, M., Hoshi, Y., Ishikawa, A., … Birbaumer, N. (2007). Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface. *NeuroImage*, *34*(4), 1416–1427. https://doi.org/10.1016/j.neuroimage.2006.11.005

Sprenger, C., Eippert, F., Finsterbusch, J., Bingel, U., Rose, M., & Büchel, C. (2012). Attention modulates spinal cord responses to pain. *Current Biology*, *22*(11), 1019–1022. https://doi.org/10.1016/j.cub.2012.04.006

Szegletes, L., Köles, M., & Forstner, B. (2014). The design of a biofeedback framework for dynamic difficulty adjustment in games. *5th IEEE International Conference on Cognitive Infocommunications, CogInfoCom 2014 - Proceedings*, 295–299. https://doi.org/10.1109/CogInfoCom.2014.7020465

Tachtsidis, I., & Scholkmann, F. (2016). False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. *Neurophotonics*, *3*(3), 039801. https://doi.org/10.1117/1.NPh.3.3.039801

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, *49*(1), 83–90. https://doi.org/10.3758/BF03211619

Thompson, M., Nordin, A. I., & Cairns, P. (2012). *Effect of Touch–Screen Size on Game Immersion*. 280–285. https://doi.org/10.14236/ewic/hci2012.38

Tiemann, L., May, E. S., Postorino, M., Schulz, E., Nickel, M. M., Bingel, U., & Ploner, M. (2015). Differential neurophysiological correlates of bottom-up and top-down modulations of pain. *Pain*, *156*(2), 289–296. https://doi.org/10.1097/01.j.pain.0000460309.94442.44

Torta, D. M., Legrain, V., Mouraux, A., & Valentini, E. (2017). Attention to pain! A neurocognitive perspective on attentional modulation of pain in neuroimaging studies. *Cortex*, *89*(January), 120–134. https://doi.org/10.1016/j.cortex.2017.01.010

Verdière, K. J., Roy, R. N., & Dehais, F. (2018). Detecting Pilot's Engagement Using fNIRS Connectivity Features in an Automated vs. Manual Landing Scenario. *Frontiers in Human Neuroscience*, *12*(January), 1–14. https://doi.org/10.3389/fnhum.2018.00006

Weiss, K. E., Dahlquist, L. M., & Wohlheiter, K. (2011). The effects of interactive and passive distraction on cold pressor pain in preschool-aged children. *Journal of Pediatric Psychology*, *36*(7), 816–826. https://doi.org/10.1093/jpepsy/jsq125

Wiener, N. (1961). Cybernetics, or control and communication in the animal and the machine, 2nd ed. In *Cybernetics, or control and communication in the animal and the machine, 2nd ed.* https://doi.org/10.1037/13140-000

Williams, A., & Ishimine, P. (2016). Non-pharmacologic Management of Pain and Anxiety in the Pediatric Patient. *Current Emergency and Hospital Medicine Reports*, *4*(1), 26–31. https://doi.org/10.1007/s40138-016-0090-5

Wohlheiter, K. A., & Dahlquist, L. M. (2013). Interactive versus passive distraction for acute pain management in young children: the role of selective attention and development. *Journal of Pediatric Psychology*, *38*(2), 202–212. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed11&NEWS=N&AN=23092971

Wong, D. L., & Baker, C. M. (1988). Pain in children: comparison of assessment scales. *The Oklahoma Nurse*, *33*(1), 8.

Wright, G. Z., Weinberger, S. J., Marti, R., & Plotzke, O. (1991). The effectiveness of infiltration anesthesia in the mandibular primary molar region. *Pediatric Dentistry*, *13*(5), 278–283.

Xiong, T., & Cherkassky, V. (2005). A combined SVM and LDA approach for classification. *Proceedings of the International Joint Conference on Neural Networks*, *3*, 1455–1459. https://doi.org/10.1109/IJCNN.2005.1556089

Yücel, M. A., Aasted, C. M., Petkov, M. P., Borsook, D., Boas, D. A., & Becerra, L. (2015). Specificity of hemodynamic brain responses to painful stimuli: a functional near-infrared spectroscopy study. *Scientific Reports*, *5*, 9469. https://doi.org/10.1038/srep09469

Zander, T. O., & Kothe, C. (2011). Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *Journal of Neural Engineering*, *8*(2), 025005. https://doi.org/10.1088/1741-2560/8/2/025005

Zander, T. O., Krol, L. R., Birbaumer, N. P., & Gramann, K. (2016). Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity. *Proceedings of the National Academy of Sciences*, *113*(52), 14898–14903. https://doi.org/10.1073/pnas.1605155114

Zhao, Y., Qiu, L., Sun, Y., Huang, C., & Li, T. (2017). Optimal hemoglobin extinction coefficient data set for near-infrared spectroscopy. *Biomedical Optics Express*, *8*(11), 5151. https://doi.org/10.1364/boe.8.005151

Zrenner, C., Belardinelli, P., Müller-Dahlhaus, F., & Ziemann, U. (2016). Closed-Loop Neuroscience and Non-Invasive Brain Stimulation: A Tale of Two Loops. *Frontiers in Cellular Neuroscience*, *10*(April), 92. https://doi.org/10.3389/fncel.2016.00092

Abdelmoniem, S. A., & Mahmoud, S. A. (2015). Comparative Evaluation of Passive, Active, and Passive-Active Distraction techniques on Pain Perception during Local Anesthesia Administration in Children. *Journal of Advanced Research*, *7*(3), 551–556. https://doi.org/10.1016/j.jare.2015.10.001

Adams, N., Little, T. D., & Ryan, R. M. (2017). Self-determination theory. *Development of Self-Determination Through the Life-Course*, *55*(1), 47–54. https://doi.org/10.1007/978-94-024-1042-6_4

Akizuki, K., & Ohashi, Y. (2015). Measurement of functional task difficulty during motor learning: What level of difficulty corresponds to the optimal challenge point? *Human Movement Science*, *43*, 107–117. https://doi.org/10.1016/j.humov.2015.07.007

Anand, K. J. S., Willson, D. F., Berger, J., Harrison, R., Meert, K. L., Zimmerman, J., … Nicholson, C. (2010). Tolerance and Withdrawal From Prolonged Opioid Use in Critically Ill Children. *Pediatrics*, *125*(5), e1208–e1225. https://doi.org/10.1542/peds.2009-0489

Andreessen, L. M., Gerjets, P., Meurers, D., & Zander, T. O. (2020). Toward neuroadaptive support technologies for improving digital reading: a passive BCI-based assessment of mental workload imposed by text difficulty and presentation speed during reading. *User Modeling and User-Adapted Interaction*. https://doi.org/10.1007/s11257-020-09273-5

Arrington, C. M., Carr, T. H., Mayer, A. R., & Rao, S. M. (2000). Neural Mechanisms of Visual Attention: Object-Based Selection of a Region in Space. *Journal of Cognitive Neuroscience*, *12*(supplement 2), 106–117. https://doi.org/10.1162/089892900563975

Auria, L., & Moro, R. A. (2011). Support Vector Machines (SVM) as a Technique for Solvency Analysis. *SSRN Electronic Journal*, (August). https://doi.org/10.2139/ssrn.1424949

Ayaz, H., Izzetoglu, M., Shewokis, P. A., & Onaral, B. (2010). Sliding-window motion artifact rejection for functional Near-Infrared Spectroscopy. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*, 6567–6570. https://doi.org/10.1109/IEMBS.2010.5627113

Ayaz, H., Onaral, B., Izzetoglu, K., Shewokis, P. A., McKendrick, R., & Parasuraman, R. (2013). Continuous monitoring of brain dynamics with functional near infrared spectroscopy as a tool for neuroergonomic research: empirical examples and a technological development. *Frontiers in Human Neuroscience*, *7*(December), 871. https://doi.org/10.3389/fnhum.2013.00871

Baker, W. B., Parthasarathy, A. B., Busch, D. R., Mesquita, R. C., Greenberg, J. H., & Yodh, A. G. (2014). Modified Beer-Lambert law for blood flow. *Biomedical Optics Express*, *5*(11), 4053. https://doi.org/10.1364/BOE.5.004053

Bantick, S. J., Wise, R. G., Ploghaus, A., Clare, S., Smith, S. M., & Tracey, I. (2002). Imaging how attention modulates pain in humans using functional MRI. *Brain : A Journal of Neurology*, *125*(Pt 2), 310–319. https://doi.org/10.1093/brain/awf022

Bastos, A. M., & Schoffelen, J. M. (2016). A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Frontiers in Systems Neuroscience*, *9*(JAN2016), 1–23. https://doi.org/10.3389/fnsys.2015.00175

Bauernfeind, G., Wriessnegger, S. C., Daly, I., & Müller-Putz, G. R. (2014). Separating heart and brain: on the reduction of physiological noise from multichannel functional near-infrared spectroscopy (fNIRS) signals. *Journal of Neural Engineering*, *11*(5), 056010. https://doi.org/10.1088/1741-2560/11/5/056010

Berger, A., Horst, F., Müller, S., Steinberg, F., & Doppelmayr, M. (2019). Current state and future prospects of EEG and fNIRS in robot-assisted gait rehabilitation: A brief review. *Frontiers in Human Neuroscience*, *13*(June). https://doi.org/10.3389/fnhum.2019.00172

Bontchev, B. (2016). Adaptation in affective video games: A literature review. *Cybernetics and Information Technologies*, *16*(3), 3–34. https://doi.org/10.1515/cait-2016-0032

Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., … Cooper, R. J. (2014). Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data. *NeuroImage*, *85*, 181–191. https://doi.org/10.1016/j.neuroimage.2013.04.082

Broad, R. D., & Wheeler, K. (2006). An adult with childhood medical trauma treated with psychoanalytic psychotherapy and EMDR: A case study. *Perspectives in Psychiatric Care*, *42*(2), 95–105. https://doi.org/10.1111/j.1744-6163.2006.00058.x

Broadbent, D. E. (1956). Successive Responses to Simultaneous Stimuli. *Quarterly Journal of Experimental Psychology*, *8*(4), 145–152. https://doi.org/10.1080/17470215608416814

Buccino, A. P., Keles, H. O., & Omurtag, A. (2016). Hybrid EEG-fNIRS asynchronous brain-computer interface for multiple motor tasks. *PLoS ONE*, *11*(1), 1–16. https://doi.org/10.1371/journal.pone.0146610

Bullinger, M., Naber, D., Pickar, D., Cohen, R. M., Kalin, N. H., Pert, A., & Bunney, W. E. (1984). Endocrine effects of the cold pressor test: Relationships to subjective pain appraisal and coping. *Psychiatry Research*, *12*(3), 227–233. https://doi.org/10.1016/0165-1781(84)90028-3

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, Vol. 97, pp. 523–547. https://doi.org/10.1037/0033-295X.97.4.523

Cairns, P., Cox, A., & Imran Nordin, A. (2014). Immersion in Digital Games: Review of Gaming Experience Research. *Handbook of Digital Games*, 339–361. https://doi.org/10.1002/9781118796443.ch12

Chanel, G., Rebetez, C., Bétrancourt, M., & Pun, T. (2008). Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. *Proceedings of the 12th International Conference on Entertainment and Media in the Ubiquitous Era - MindTrek '08*, (August), 13. https://doi.org/10.1145/1457199.1457203

Chanel, G., Rebetez, C., Bétrancourt, M., & Pun, T. (2011). *Emotion Assessment From Physiological Signals for Adaptation of Game Difficulty*. *41*(6), 1052–1063.

City, R. (2017). *(19) United States (12)* (Vol. 1).

Cooper, T. E., Fisher, E., Gray, A. L., Krane, E., Sethna, N., van Tilburg, M. A. L., … Wiffen, P. J. (2017). Opioids for chronic non-cancer pain in children and adolescents. *Cochrane Database of*

Systematic Reviews, *2017*(2). https://doi.org/10.1002/14651858.CD012538

Cope, M. (1991). The development of a near infrared spectroscopy system and its application for non invasive monitory of cerebral blood and tissue oxygenation in the newborn infants. *Diss. University of London*, 342. Retrieved from http://discovery.ucl.ac.uk/1317956/

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Erratum to "Voluntary orienting is dissociated from target detection in human posterior parietal cortex." *Nature Neuroscience*, *3*(5), 521–521. https://doi.org/10.1038/74905

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Coyle, S. M., Ward, T. E., Markham, C. M., Ward, E., & Markham, C. M. (2007). Brain-computer interface using a simplified functional near-infrared spectroscopy system. *Journal of Neural Engineering*, *4*(3), 219–226. https://doi.org/10.1088/1741-2560/4/3/007

Crombez, G., Van Ryckeghem, D. M. L., Eccleston, C., & Van Damme, S. (2013). Attentional bias to pain-related information: A meta-analysis. *Pain*, *154*(4), 497–510. https://doi.org/10.1016/j.pain.2012.11.013

Csikszentmihalyi, M. (1975). Beyond Boredom and Anxiety: Experiencing Flow in Work and Play. *The Jossey-Bass Behavioral Science Series*, 231. https://doi.org/10.2307/2065805

Cui, X., Bray, S., & Reiss, A. L. (2010a). Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *NeuroImage*, *49*(4), 3039–3046. https://doi.org/10.1016/j.neuroimage.2009.11.050

Cui, X., Bray, S., & Reiss, A. L. (2010b). Speeded near infrared spectroscopy (NIRS) response detection. *PLoS ONE*, *5*(11). https://doi.org/10.1371/journal.pone.0015474

Cui, X., Bryant, D. M., & Reiss, A. L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *NeuroImage*, *59*(3), 2430–2437. https://doi.org/10.1016/j.neuroimage.2011.09.003

Czub, M., & Piskorz, J. (2014). How body movement influences virtual reality analgesia? *Proceedings - 2014 International Conference on Interactive Technologies and Games, ITAG 2014*, 13–19. https://doi.org/10.1109/iTAG.2014.8

Dehais, F., Causse, M., Vachon, F., Régis, N., Menant, E., & Tremblay, S. (2014). Failure to detect critical auditory alerts in the cockpit: Evidence for inattentional deafness. In *Human Factors* (Vol. 56). https://doi.org/10.1177/0018720813510735

Diseth, T. H. (2006). Dissociation following traumatic medical treatment procedures in childhood: A longitudinal follow-up. *Development and Psychopathology*, *18*(1), 233–251. https://doi.org/10.1017/S0954579406060135

Durantin, G., Dehais, F., Gonthier, N., Terzibas, C., & Callan, D. E. (2017). Neural signature of inattentional deafness. *Human Brain Mapping*, *38*(11), 5440–5455. https://doi.org/10.1002/hbm.23735

Ebrahimpour, F., Sadeghi, N., Najafi, M., Iraj, B., & Shahrokhi, A. (2015). Effect of Playing Interactive Computer Game on Distress of Insulin Injection Among Type 1 Diabetic Children. *Iranian*

Journal of Pediatrics, *25*(3), e427. https://doi.org/10.5812/ijp.25(3)2015.427

Eccleston, C., & Crombez, G. (1999). Pain demands attention: a cognitive-affective model of the interruptive function of pain. *Psychol Bull*, *125*(3), 356–366. https://doi.org/10.1037/0033-2909.125.3.356

Ewing, K. C., Fairclough, S. H., & Gilleade, K. (2016). Evaluation of an Adaptive Game that Uses EEG Measures Validated during the Design Process as Inputs to a Biocybernetic Loop. *Frontiers in Human Neuroscience*, *10*(May), 1–13. https://doi.org/10.3389/fnhum.2016.00223

Fairclough, S., & Gilleade, K. (2012a). Construction of the Biocybernetic Loop: A Case Study. *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, *Santa Moni*, 571–578. https://doi.org/10.1145/2388676.2388797

Fairclough, S., & Gilleade, K. (2012b). Construction of the biocybernetic loop. *Proceedings of the 14th ACM International Conference on Multimodal Interaction - ICMI '12*, 571. https://doi.org/10.1145/2388676.2388797

Fairclough, S. H. (2009a). Fundamentals of physiological computing. *Interacting with Computers*, *21*(1–2), 133–145. https://doi.org/10.1016/j.intcom.2008.10.011

Fairclough, S. H. (2009b). Fundamentals of physiological computing. *Interacting with Computers*, *21*(1–2), 133–145. https://doi.org/10.1016/j.intcom.2008.10.011

Fairclough, S. H. (2010). *Physiological Computing: Interfacing with the Human Nervous System*. 1–20. https://doi.org/10.1007/978-90-481-3258-4_1

Fairclough, S. H., Gilleade, K., Ewing, K. C., & Roberts, J. (2013a). Capturing user engagement via psychophysiology: measures and mechanisms for biocybernetic adaptation. *International Journal of Autonomous and Adaptive Communications Systems*, *6*(1), 63. https://doi.org/10.1504/IJAACS.2013.050694

Fairclough, S. H., Gilleade, K., Ewing, K. C., & Roberts, J. (2013b). Capturing user engagement via psychophysiology: Measures and mechanisms for biocybernetic adaptation. *International Journal of Autonomous and Adaptive Communications Systems*, *6*(1), 63–79. https://doi.org/10.1504/IJAACS.2013.050694

Fairclough, S. H., Stamp, K., Dobbins, C., & Poole, H. M. (2020). Computer games as distraction from PAIN: Effects of hardware and difficulty on pain tolerance and subjective IMMERSION. *International Journal of Human Computer Studies*, *139*(May 2019). https://doi.org/10.1016/j.ijhcs.2020.102427

Fernández, H. D., Koji, M., & Kondo, K. (2017). Adaptable game experience based on player's performance and EEG. *Proceedings - 2017 NICOGRAPH International, NICOInt 2017*, (June), 1–8. https://doi.org/10.1109/NICOInt.2017.11

Ferrari, M., & Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage*, *63*(2), 921–935. https://doi.org/10.1016/j.neuroimage.2012.03.049

Fishburn, F. a, Norr, M. E., Medvedev, A. V, & Vaidya, C. J. (2014). Sensitivity of fNIRS to cognitive state and load. *Frontiers in Human Neuroscience*, *8*(February), 76. https://doi.org/10.3389/fnhum.2014.00076

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(26), 10046–10051. https://doi.org/10.1073/pnas.0604187103

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(27), 9673–9678. https://doi.org/10.1073/pnas.0504136102

Freeman, F. G., Mikulka, P. J., Prinzel, L. J., & Scerbo, M. W. (1999). Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology*, *50*(1), 61–76. https://doi.org/10.1016/S0301-0511(99)00002-2

Girouard, A., Solovey, E. T., Hirshfield, L. M., Chauncey, K., Sassaroli, A., Fantini, S., & Jacob, R. J. K. (2009). Distinguishing Difficulty Levels with Non-invasive Brain. *Proceedings of INTERACT 2009*, (Figure 1), 440–452.

Goldman, A., & Lloyd-Thomas, A. R. (2004). Pain management in children. *Current Anaesthesia and Critical Care*, *15*(4–5), 272–283. https://doi.org/10.1016/j.cacc.2004.08.014

Goodwin, J. R., Gaudet, C. R., & Berger, A. J. (2014). Short-channel functional near-infrared spectroscopy regressions improve when source-detector separation is reduced. *Neurophotonics*, *1*(1), 015002. https://doi.org/10.1117/1.nph.1.1.015002

Gordon, N. S., Merchant, J., Zanbaka, C., Hodges, L. F., & Goolkasian, P. (2011). Interactive gaming reduces experimental pain with or without a head mounted display. *Computers in Human Behavior*, *27*(6), 2123–2128. https://doi.org/10.1016/j.chb.2011.06.006

Gould, D., Goldstone, L., Kelly, D., & Gammon, J. (2004). Examining the validity of pressure ulcer risk assessment scales: A replication study. *International Journal of Nursing Studies*, *41*(3), 331–339. https://doi.org/10.1016/j.ijnurstu.2003.10.005

Gouraud, J., Delorme, A., & Berberian, B. (2017). Autopilot, mind wandering, and the out of the loop performance problem. *Frontiers in Neuroscience*, *11*(OCT), 1–13. https://doi.org/10.3389/fnins.2017.00541

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, *11*(5/6), 561–566. https://doi.org/10.5194/npg-11-561-2004

Guadagnoll, M. A., & Lee, T. D. (2004). Challenge Point: A Framework for Conceptualizing the Effects of Various Practice Conditions in Motor Learning. *Journal of Motor Behavior*, *36*(2), 212–224. https://doi.org/10.3200/JMBR.36.2.212-224

Gupta, D., Agarwal, A., Dhiraaj, S., Tandon, M., Kumar, M., Singh, R. S., … Singh, U. (2006). An evaluation of efficacy of balloon inflation on venous cannulation pain in children: A prospective, randomized, controlled study. *Anesthesia and Analgesia*, *102*(5), 1372–1375. https://doi.org/10.1213/01.ane.0000205741.82299.d6

Gwinup, G., Haw, T., & Elias, A. (1983). Cardiovascular changes in video-game players. *Postgraduate Medicine*, *74*(6), 245–248. https://doi.org/10.1080/00325481.1983.11698546

Harrivel, A. R., Weissman, D. H., Noll, D. C., & Peltier, S. J. (2013). Monitoring attentional state with fNIRS. *Frontiers in Human Neuroscience*, *7*(December), 861. https://doi.org/10.3389/fnhum.2013.00861

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, *52*, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Heger, D., Herff, C., Putze, F., Mutter, R., & Schultz, T. (2014). Continuous Affective States Recognition using functional Near Infrared Spectroscopy. *Brain Computer Interfaces*. https://doi.org/DOI 10.1080/2326263X.2014.912884

Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F., & Schultz, T. (2014). Mental workload during n-back task—quantified in the prefrontal cortex using fNIRS. *Frontiers in Human Neuroscience*, *7*(January), 1–9. https://doi.org/10.3389/fnhum.2013.00935

Hirshfield, L. M., Chauncey, K., Gulotta, R., Girouard, A., Solovey, E. T., Jacob, R. J. K., … Fantini, S. (2009). Combining Electroencephalograph and Near Infrared Spectroscopy to Explore Users' Instantaneous and Continuous Mental Workload States. *HCI International 2009 13th International Conference on Human-Computer Interaction*, 239–247. https://doi.org/10.1007/978-3-642-02812-0_28

Hoffman, H. G. (2004). Virtual-reality therapy. *Scientific American*, *291*(2), 58–65. https://doi.org/10.1038/scientificamerican0804-58

Hoffman, H. G., Doctor, J. N., Patterson, D. R., Carrougher, G. J., & Furness, T. A. (2000). Virtual reality as an adjunctive pain control during burn wound care in adolescent patients. *Pain*, *85*(1–2), 305–309. https://doi.org/10.1016/S0304-3959(99)00275-4

Hou, J., Nam, Y., Peng, W., & Lee, K. M. (2012). Effects of screen size, viewing angle, and players' immersion tendencies on game experience. *Computers in Human Behavior*, *28*(2), 617–623. https://doi.org/10.1016/j.chb.2011.11.007

Hunicke, R. (2005). The case for dynamic difficulty adjustment in games. *ACM International Conference Proceeding Series*, *265*(January 2005), 429–433. https://doi.org/10.1145/1178477.1178573

Hussein, H. A. (2015). Effect of active and passive distraction on decreasing pain associated with painful medical procedures among school aged children. *World J Nurs Sci*, *1*(2), 13–23. https://doi.org/10.5829/idosi.wjns.2015.1.2.93202

Inan, G., & Inal, S. (2019). The impact of 3 different distraction techniques on the pain and anxiety levels of children during venipuncture. *Clinical Journal of Pain*, *35*(2), 140–147. https://doi.org/10.1097/AJP.0000000000000666

Jameson, E., Trevena, J., & Swain, N. (2011a). Electronic gaming as pain distraction. *Pain Research and Management*, *16*(1), 27–32. https://doi.org/10.1155/2011/856014

Jameson, E., Trevena, J., & Swain, N. (2011b). Electronic gaming as pain distraction. *Pain Research and Management*, *16*(1), 27–32.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human Computer*

Studies, *66*(9), 641–661. https://doi.org/10.1016/j.ijhcs.2008.04.004

Jennings-Teats, M., Smith, G., & Wardrip-Fruin, N. (2010). Polymorph: Dynamic Difficulty Adjustment through level generation. *Workshop on Procedural Content Generation in Games, PC Games 2010, Co-Located with the 5th International Conference on the Foundations of Digital Games*, (September), 2–6. https://doi.org/10.1145/1814256.1814267

Johnson, M. H. (2005). How does distraction work in the management of pain? *Current Pain and Headache Reports*, *9*(2), 90–95. https://doi.org/10.1007/s11916-005-0044-1

Julien, C. (2006). The enigma of Mayer waves: Facts and models. *Cardiovascular Research*, *70*(1), 12–21. https://doi.org/10.1016/j.cardiores.2005.11.008

Kaheni, S., Rezai, M. S., Bagheri-nesami, M., & Hossein, A. (n.d.). *The Effect of Distraction Technique on the Pain of Dressing Change among 3-6 Year-old Children*. *4*(28), 1603–1610.

Kamran, M. A., Mannann, M. M. N., & Jeong, M. Y. (2018). Differential path-length factor's effect on the characterization of brain's hemodynamic response function: A functional near-infrared study. *Frontiers in Neuroinformatics*, *12*(June), 1–15. https://doi.org/10.3389/fninf.2018.00037

Keogh, E., Moore, D. J., Duggan, G. B., Payne, S. J., & Eccleston, C. (2013). The disruptive effects of pain on complex cognitive performance and executive control. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0083272

Kincade, J. M., Abrams, R. A., Astafiev, S. V., Shulman, G. L., & Corbetta, M. (2005). An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *Journal of Neuroscience*, *25*(18), 4593–4604. https://doi.org/10.1523/JNEUROSCI.0236-05.2005

Koenraadt, K. L. M., Roelofsen, E. G. J., Duysens, J., & Keijsers, N. L. W. (2014). Cortical control of normal gait and precision stepping: An fNIRS study. *NeuroImage*, *85*, 415–422. https://doi.org/10.1016/j.neuroimage.2013.04.070

Kononenko, I., Šimec, E., & Robnik-Šikonja, M. (1997). Overcoming the myopia of inductive learning algorithms with RELIEFF. *Applied Intelligence*, *7*(1), 39–55. https://doi.org/10.1023/A:1008280620621

Krol, L. R., Freytag, S. C., & Zander, T. O. (2017). Meyendtris: A hands-free, multimodal tetris clone using eye tracking and passive BCI for intuitive neuroadaptive gaming. *ICMI 2017 - Proceedings of the 19th ACM International Conference on Multimodal Interaction*, *2017-Janua*, 433–437. https://doi.org/10.1145/3136755.3136805

Laird, A. R., Eickhoff, S. B., Rottschy, C., Bzdok, D., Ray, K. L., & Fox, P. T. (2013). Networks of task co-activations. *NeuroImage*, *80*, 505–514. https://doi.org/10.1016/j.neuroimage.2013.04.073

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N., Beck, D. M., & Konstantinou, N. (2014). Blinded by the load: Attention, awareness and the role of perceptual load. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1641). https://doi.org/10.1098/rstb.2013.0205

Law, Dahlquist, Still, Weiss, Herbert, Wohlheiter, & S, H. (2011). Videogame distraction using virtual reality technology for children exeperiencing cold pressor pain: the role of cognitive processing.

Journal of Pediatric Psychology, *36*(1), 84–94.

Leamy, D. J., Collins, R., & Ward, T. E. (2011). Combining fNIRS and EEG to improve motor cortex activity classification during an imagined movement-based task. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6780 LNAI*, 177–185. https://doi.org/10.1007/978-3-642-21852-1_22

Legrain, V., Damme, S. Van, Eccleston, C., Davis, K. D., Seminowicz, D. A., & Crombez, G. (2009). A neurocognitive model of attention to pain: Behavioral and neuroimaging evidence. *Pain*, *144*(3), 230–232. https://doi.org/10.1016/j.pain.2009.03.020

Liu, C., Agrawal, P., Sarkar, N., & Chen, S. (2009). Dynamic Difficulty Adjustment in Computer Games Through Real-Time Anxiety-Based Affective Feedback. *International Journal of Human-Computer Interaction*, *25*(6), 506–529. https://doi.org/10.1080/10447310902963944

Liu, Y., Piazza, E. A., Simony, E., Shewokis, P. A., Onaral, B., Hasson, U., & Ayaz, H. (2017). Measuring speaker-listener neural coupling with functional near infrared spectroscopy. *Scientific Reports*, *7*(February), 1–13. https://doi.org/10.1038/srep43293

Lovallo, W. (1975). The Cold Pressor Test and Autonomic Function: A Review and Integration. *Psychophysiology*, *12*(3), 268–282. https://doi.org/10.1111/j.1469-8986.1975.tb01289.x

MacLaren, J. E., & Cohen, L. L. (2005). A comparison of distraction strategies for venipuncture distress in children. *Journal of Pediatric Psychology*, *30*(5), 387–396. https://doi.org/10.1093/jpepsy/jsi062

Matson, K. L., Johnson, P. N., Tran, V., Horton, E. R., & Sterner-Allison, J. (2019). Opioid use in children. *Journal of Pediatric Pharmacology and Therapeutics*, *24*(1), 72–75. https://doi.org/10.5863/1551-6776-24.1.72

Matthews, G., Joyner, L., Gilliand, K., Campbell, S., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire- Towards a state “big three”? *Personality Psychology*, *7*, 335–350.

Mayer, A. R., Toulouse, T., Klimaj, S., Ling, J. M., Pena, A., & Bellgowan, P. S. F. (2014). Investigating the Properties of the Hemodynamic Response Function after Mild Traumatic Brain Injury. *Journal of Neurotrauma*, *31*(2), 189–197. https://doi.org/10.1089/neu.2013.3069

Meng, L., Pei, G., Zheng, J., & Ma, Q. (2016). Close games versus blowouts: Optimal challenge reinforces one’s intrinsic motivation to win. *International Journal of Psychophysiology*, *110*, 102–108. https://doi.org/10.1016/j.ijpsycho.2016.11.001

Metz, A. J., Wolf, M., Achermann, P., & Scholkmann, F. (2015). A new approach for automatic removal of movement artifacts in near-infrared spectroscopy time series by means of acceleration data. *Algorithms*, *8*(4), 1052–1075. https://doi.org/10.3390/a8041052

Miller, K., Rodger, S., Bucolo, S., Greer, R., & Kimble, R. M. (2010). Multi-modal distraction. Using technology to combat pain in young children with burn injuries. *Burns*, *36*(5), 647–658. https://doi.org/10.1016/j.burns.2009.06.199

Mitchell, L. A., MacDonald, R. A. R., & Brodie, E. E. (2004). Temperature and the cold pressor test. *The Journal of Pain*, *5*(4), 233–237. https://doi.org/10.1016/j.jpain.2004.03.004

Muñoz, J. E., Gouveia, E. R., Cameirão, M. S., & Badia, S. B. I. (2017). The biocybernetic loop engine: An integrated tool for creating physiologically adaptive videogames. *PhyCS 2017 - Proceedings of the 4th International Conference on Physiological Computing Systems*, (PhyCS), 45–54. https://doi.org/10.5220/0006429800450054

Nacke, L. E., & Lindley, C. A. (2008). Flow and Immersion in First-person Shooters: Measuring the Player's Gameplay Experience. *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, 81–88. https://doi.org/10.1145/1496984.1496998

Naito, M., Michioka, Y., Ozawa, K., Ito, Y. I., Kiguchi, M., & Kanazawa, T. (2007). A communication means for totally locked-in ALS patients based on changes in cerebral blood volume measured with near-infrared light. *IEICE Transactions on Information and Systems*, *E90-D*(7), 1028–1037. https://doi.org/10.1093/ietisy/e90-d.7.1028

Naseer, N., & Hong, K.-S. (2015). fNIRS-based brain-computer interfaces: a review. *Frontiers in Human Neuroscience*, *9*(January), 1–15. https://doi.org/10.3389/fnhum.2015.00003

Nilsson, S., Enskär, K., Hallqvist, C., & Kokinsky, E. (2013). Active and Passive Distraction in Children Undergoing Wound Dressings. *Journal of Pediatric Nursing*, *28*(2), 158–166. https://doi.org/10.1016/j.pedn.2012.06.003

Norman, D. A., & Shallice, T. (1980). *Attention to action: Willed and automatic control of behavior. Report No 8006.* (January 1986), 43. https://doi.org/10.1007/978-1-4757-0629-1_1

Nouwen, A., Cloutier, C., Kappas, A., Warbrick, T., & Sheffield, D. (2006). Effects of Focusing and Distraction on Cold Pressor–Induced Pain in Chronic Back Pain Patients and Control Subjects. *The Journal of Pain*, *7*(1), 62–71. https://doi.org/10.1016/j.jpain.2005.08.004

Parker, M., Delahunty, B., Heberlein, N., Devenish, N., Wood, F. M., Jackson, T., … Edgar, D. W. (2016). Interactive gaming consoles reduced pain during acute minor burn rehabilitation: A randomized, pilot trial. *Burns*, *42*(1), 91–96. https://doi.org/10.1016/j.burns.2015.06.022

Pfeifer, M. D., Scholkmann, F., & Labruyère, R. (2018). Signal Processing in Functional Near-Infrared Spectroscopy (fNIRS): Methodological Differences Lead to Different Statistical Results. *Frontiers in Human Neuroscience*, *11*(January), 1–12. https://doi.org/10.3389/fnhum.2017.00641

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, *40*(1–2), 187–195. https://doi.org/10.1016/0301-0511(95)05116-3

Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2000). A Closed-Loop System for Examining Psychophysiological Measures for Adaptive Task Allocation. *The International Journal of Aviation Psychology*, *10*(4), 393–410. https://doi.org/10.1207/S15327108IJAP1004_6

Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2003). Effects of a Psychophysiological System for Adaptive Automation on Performance, Workload, and the Event-Related Potential P300 Component. *Human Factors*, *45*(4), 601–613. https://doi.org/10.1518/hfes.45.4.601.27092

Przybylski, A. K., Deci, E. L., Rigby, C. S., & Ryan, R. M. (2014). Competence-Impeding Electronic Games and Players' Aggressive Feelings, Thoughts, and Behaviors. *Journal of Personality and Social Psychology*, *106*(3), 441–457. https://doi.org/10.1037/a0034820

Racz, F. S., Mukli, P., Nagy, Z., & Eke, A. (2017). Increased prefrontal cortex connectivity during cognitive challenge assessed by fNIRS imaging. *Biomedical Optics Express*, *8*(8), 3842. https://doi.org/10.1364/boe.8.003842

Rémy, F., Frankenstein, U. N., Mincic, A., Tomanek, B., & Stroman, P. W. (2003). Pain modulates cerebral activity during cognitive performance. *NeuroImage*, *19*(3), 655–664. https://doi.org/10.1016/S1053-8119(03)00146-0

Rheinberg, F. (2008). Intrinsic Motivation and Flow. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation and Action* (pp. 323–348). https://doi.org/10.1017/CBO9780511499821.014

Rutter, C. E., Dahlquist, L. M., & Weiss, K. E. (2009). Sustained Efficacy of Virtual Reality Distraction. *The Journal of Pain*, *10*(4), 391–397. https://doi.org/10.1016/j.jpain.2008.09.016

Scalf, P. E., Torralbo, A., Tapia, E., & Beck, D. M. (2013). Competition explains limited attention and perceptual resources: Implications for perceptual load and dilution theories. *Frontiers in Psychology*, *4*(MAY), 1–9. https://doi.org/10.3389/fpsyg.2013.00243

Scerbo, M. W., Freeman, F. G., & Mikulka, P. J. (2003). A brain-based system for adaptive automation. *Theoretical Issues in Ergonomics Science*, *4*(1–2), 200–219. https://doi.org/10.1080/1463922021000020891

Scholkmann, F., & Wolf, M. (2013). General equation for the differential pathlength factor of the frontal human head depending on wavelength and age. *Journal of Biomedical Optics*, *18*(10), 105004. https://doi.org/10.1117/1.jbo.18.10.105004

Selesnick, I. W., & Burrus, C. S. (1998). *H 1 = h 2 = a=d . 46*(6), 1688–1694.

Sitaram, R., Zhang, H., Guan, C., Thulasidas, M., Hoshi, Y., Ishikawa, A., … Birbaumer, N. (2007). Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface. *NeuroImage*, *34*(4), 1416–1427. https://doi.org/10.1016/j.neuroimage.2006.11.005

Sprenger, C., Eippert, F., Finsterbusch, J., Bingel, U., Rose, M., & Büchel, C. (2012). Attention modulates spinal cord responses to pain. *Current Biology*, *22*(11), 1019–1022. https://doi.org/10.1016/j.cub.2012.04.006

Szegletes, L., Köles, M., & Forstner, B. (2014). The design of a biofeedback framework for dynamic difficulty adjustment in games. *5th IEEE International Conference on Cognitive Infocommunications, CogInfoCom 2014 - Proceedings*, 295–299. https://doi.org/10.1109/CogInfoCom.2014.7020465

Tachtsidis, I., & Scholkmann, F. (2016). False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. *Neurophotonics*, *3*(3), 039801. https://doi.org/10.1117/1.NPh.3.3.039801

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, *49*(1), 83–90. https://doi.org/10.3758/BF03211619

Thompson, M., Nordin, A. I., & Cairns, P. (2012). *Effect of Touch–Screen Size on Game Immersion*. 280–285. https://doi.org/10.14236/ewic/hci2012.38

Tiemann, L., May, E. S., Postorino, M., Schulz, E., Nickel, M. M., Bingel, U., & Ploner, M. (2015). Differential neurophysiological correlates of bottom-up and top-down modulations of pain. *Pain*, *156*(2), 289–296. https://doi.org/10.1097/01.j.pain.0000460309.94442.44

Torta, D. M., Legrain, V., Mouraux, A., & Valentini, E. (2017). Attention to pain! A neurocognitive perspective on attentional modulation of pain in neuroimaging studies. *Cortex*, *89*(January), 120–134. https://doi.org/10.1016/j.cortex.2017.01.010

Verdière, K. J., Roy, R. N., & Dehais, F. (2018). Detecting Pilot's Engagement Using fNIRS Connectivity Features in an Automated vs. Manual Landing Scenario. *Frontiers in Human Neuroscience*, *12*(January), 1–14. https://doi.org/10.3389/fnhum.2018.00006

Weiss, K. E., Dahlquist, L. M., & Wohlheiter, K. (2011). The effects of interactive and passive distraction on cold pressor pain in preschool-aged children. *Journal of Pediatric Psychology*, *36*(7), 816–826. https://doi.org/10.1093/jpepsy/jsq125

Wiener, N. (1961). Cybernetics, or control and communication in the animal and the machine, 2nd ed. In *Cybernetics, or control and communication in the animal and the machine, 2nd ed.* https://doi.org/10.1037/13140-000

Williams, A., & Ishimine, P. (2016). Non-pharmacologic Management of Pain and Anxiety in the Pediatric Patient. *Current Emergency and Hospital Medicine Reports*, *4*(1), 26–31. https://doi.org/10.1007/s40138-016-0090-5

Wohlheiter, K. A., & Dahlquist, L. M. (2013). Interactive versus passive distraction for acute pain management in young children: the role of selective attention and development. *Journal of Pediatric Psychology*, *38*(2), 202–212. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed11&NEWS=N&AN=23092971

Wong, D. L., & Baker, C. M. (1988). Pain in children: comparison of assessment scales. *The Oklahoma Nurse*, *33*(1), 8.

Wright, G. Z., Weinberger, S. J., Marti, R., & Plotzke, O. (1991). The effectiveness of infiltration anesthesia in the mandibular primary molar region. *Pediatric Dentistry*, *13*(5), 278–283.

Xiong, T., & Cherkassky, V. (2005). A combined SVM and LDA approach for classification. *Proceedings of the International Joint Conference on Neural Networks*, *3*, 1455–1459. https://doi.org/10.1109/IJCNN.2005.1556089

Yücel, M. A., Aasted, C. M., Petkov, M. P., Borsook, D., Boas, D. A., & Becerra, L. (2015). Specificity of hemodynamic brain responses to painful stimuli: a functional near-infrared spectroscopy study. *Scientific Reports*, *5*, 9469. https://doi.org/10.1038/srep09469

Zander, T. O., & Kothe, C. (2011). Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *Journal of Neural Engineering*, *8*(2), 025005. https://doi.org/10.1088/1741-2560/8/2/025005

Zander, T. O., Krol, L. R., Birbaumer, N. P., & Gramann, K. (2016). Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity. *Proceedings of the National Academy of Sciences*, *113*(52), 14898–14903. https://doi.org/10.1073/pnas.1605155114

Zhao, Y., Qiu, L., Sun, Y., Huang, C., & Li, T. (2017). Optimal hemoglobin extinction coefficient data set for near-infrared spectroscopy. *Biomedical Optics Express*, *8*(11), 5151.

https://doi.org/10.1364/boe.8.005151

Zrenner, C., Belardinelli, P., Müller-Dahlhaus, F., & Ziemann, U. (2016). Closed-Loop Neuroscience and Non-Invasive Brain Stimulation: A Tale of Two Loops. *Frontiers in Cellular Neuroscience*, *10*(April), 92. https://doi.org/10.3389/fncel.2016.00092

Appendix 1: The Immersive Experience Questionnaire

Immersion Questionnaire used in Experiment 3

Your Experience of the Game

Please answer the following questions by circling the relevant number. In particular, remember that these questions are asking you about how you felt at the end of the game.

To what extent did the game hold your attention?

Not at all 1 2 3 4 5 A lot

To what extent did you feel you were focused on the game?

Not at all 1 2 3 4 5 A lot

How much effort did you put into playing the game?

Very little 1 2 3 4 5 A lot

Did you feel that you were trying you best?

Not at all 1 2 3 4 5 Very much so

To what extent did you lose track of time?

Not at all 1 2 3 4 5 A lot

To what extent did you feel consciously aware of being in the real world whilst playing?

Not at all 1 2 3 4 5 Very much so

To what extent did you forget about your everyday concerns?

Not at all 1 2 3 4 5 A lot

To what extent were you aware of yourself in your surroundings?

Not at all 1 2 3 4 5 Very Aware

To what extent did you notice events taking place around you?

Not at all 1 2 3 4 5 A lot

Did you feel the urge at any point to stop playing and see what was happening around you?

Not at all 1 2 3 4 5 Very much so

To what extent did you feel that you were interacting with the game environment?

Not at all 1 2 3 4 5 Very much so

To what extent did you feel as though you were separated from your real-world environment?

Not at all 1 2 3 4 5 Very much so

To what extent did you feel that the game was something you were experiencing, rather than something you were just doing?

Not at all 1 2 3 4 5 Very much so

To what extent was your sense of being in the game environment stronger than your sense of being in the real world?

Not at all 1 2 3 4 5 Very much so

At any point did you find yourself become so involved that you were unaware you were even using controls?

Not at all 1 2 3 4 5 Very much so

To what extent did you feel as though you were moving through the game according to you own will?

Not at all 1 2 3 4 5 Very much so

To what extent did you find the game challenging?

Not at all 1 2 3 4 5 Very difficult

Were there any times during the game in which you just wanted to give up?

Not at all 1 2 3 4 5 A lot

To what extent did you feel motivated while playing?

Not at all 1 2 3 4 5 A lot

To what extent did you find the game easy?

Not at all 1 2 3 4 5 Very much so

To what extent did you feel like you were making progress towards the end of the game?

Not at all 1 2 3 4 5 A lot

How well do you think you performed in the game?

Very Poor 1 2 3 4 5 Very well

To what extent did you feel emotionally attached to the game?

Not at all 1 2 3 4 5 Very much so

To what extent were you interested in seeing how the game's events would progress?

Not at all 1 2 3 4 5 A lot

How much did you want to "win" the game?

Not at all 1 2 3 4 5 Very much so

Were you in suspense about whether or not you would win or lose the game?

Not at all 1 2 3 4 5 Very much so

At any point did you find yourself become so involved that you wanted to speak to the game directly?

Not at all 1 2 3 4 5 Very much so

To what extent did you enjoy the graphics and the imagery?

Not at all 1 2 3 4 5 A lot

How much would you say you enjoyed playing the game?

Not at all 1 2 3 4 5A lot

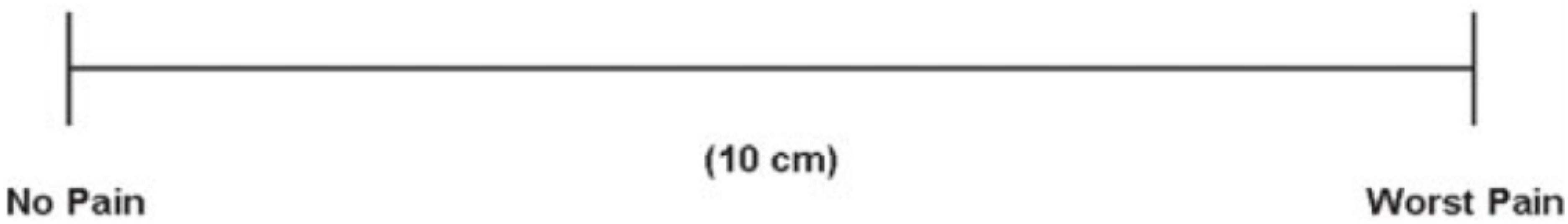
When interrupted, were you disappointed that the game was over?

Not at all 1 2 3 4 5 Very much so

Would you like to play the game again?

Definitely not 1 2 3 4 5 Definitely yes

Appendix 2: Visual Analogue Scale



Appendix 3: National Aeronautics and Space Administration Task Load Index (NASA TLX)

Please rate the MENTAL DEMAND of the game: How much mental and perceptual activity was required

Low										High
0	1	2	3	4	5	6	7	8	9	10

Enter a number between 1 and 10 here for MENTAL DEMAND

Please rate the PHYSICAL DEMAND of the game: How much physical activity was required?

Low										High
0	1	2	3	4	5	6	7	8	9	10

Enter a number between 1 and 10 here for PHYSICAL DEMAND

Please rate the TEMPORAL DEMAND of the game: How much time pressure did you feel due to the pace at which task elements occurred?

Low										High
0	1	2	3	4	5	6	7	8	9	10

Enter a number between 1 and 10 here for TEMPORAL DEMAND

Please rate your own PERFORMANCE during the game: How successful do you think you were?

Low High

0	1	2	3	4	5	6	7	8	9	10

Enter a number between 1 and 10 here for PERFORMANCE

Please rate your EFFORT: How hard did you have to work mentally to accomplish your level of performance during the game?

Low										High
0	1	2	3	4	5	6	7	8	9	10

Enter a number between 1 and 10 here for EFFORT

Please rate your FRUSTRATION: How discouraged, irritated, stressed and annoyed did you feel during the game?

Low										High
0	1	2	3	4	5	6	7	8	9	10

Enter a number between 1 and 10 here for FRUSTRATION

Appendix 4: Motivation Scale

MOTIVATION

Please answer some questions about your attitude to the task you have just done. Rate your agreement with the following statements by typing in the most appropriate number.

0 = Not At All
1 = A Little Bit
2 = Somewhat
3 = Very Much
4 = Extremely

		ANSWERS	
1	I felt motivated when doing the game		
2	The content of the game was interesting.		
3	I did not care about doing well on the game.		
4	I feel annoyed after having completed the game.		
5	I exerted a great deal of effort into the game.		
6	I did not care about suceeding on the game.		
7	Performing badly on the game will make me feel upset.		
8	Doing the game was a waste of time.		